DOCUMENT RESUME

ED 316 066 HE 022 211

AUTHOR	Blackburn, Robert T.; Pitney, Judith A.
TITLE	Performance Appraisal for Faculty. Implications for Higher Education. From the Program on Faculty as a Key Resource.
INSTITUTION	National Center for Research to Improve Postsecondary Teaching and Learning, Ann Arbor, MI.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO	NCRIPTAL-TR-88-D-002.0
PUB DATE	88
GRANT	OERI-86-0010
NOTE	59p.
AVAILABLE FROM	The National Center for Research to Improve Postsecondary Teaching and Learning, 2400 School of Education Bldg., The University of Michigan, Ann Arbor, MI 48109 ($10.00).
PUB TYPE	Information Analyses (070)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Change Strategies; *College Faculty; College Students; Educational Assessment; *Educational Quality; Excellence in Education; Faculty Development; Higher Education; *Instructional Improvement; Learning; Literature Reviews; *Personnel Evaluation; Student Development; *Teacher Evaluation; Teacher Improvement

ABSTRACT

This review of the literature on performance appraisal for college and university faculty was undertaken to serve the ultimate goal of improving student cognitive learning. The starting point was to ask what is known about performance appraisal. What are its positive and negative consequences? How does the manner in which it is conducted relate to future performance? In what ways do the skills of the appraiser affect the outcomes? To what degree can what is accepted about performance appraisal in other settings be expected to be true for faculty and department chairs in colleges and universities? These and other questions directed the investigation. The review contains five sections and begins with a brief exposition of the principal theories that underlie performance appraisal research in Section I. Section II examines the research studies conducted in the field or in the laboratory to test hypotheses drawn from the theories. The research is principally in the psychological and organizational fields with few contributions from postsecondary education. Section II ends with 15 general conclusions, "truths" that postsecondary education needs to take into consideration to improve its performance and more closely achieve its goals. Section III describes the norms, values, structures, and practices that characterize colleges and universities so that the application of the findings can be inferred as to their likelihood of successful outcomes in postsecondary institutions. Following this, a plan for performance appraisal well suited for colleges and universities is advanced in Section IV and the conditions under which it could be successful are discussed in Section V. The list of cited references, an appendix, and a bibliography complete the monograph. The appendix briefly identifies the main appraisal procedures currently in practice in business. (AA)

ED316066

NATIONAL CENTER FOR
RESEARCH TO IMPROVE
POSTSECONDARY
TEACHING AND LEARNING

Performance Appraisal for Faculty: Implications for Higher Education

by
Robert T. Blackburn
and Judith A. Pitney

FROM THE PROGRAM ON FACULTY AS A KEY RESOURCE

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

NCRIPTAL

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

022 211

Performance Appraisal for Faculty

Implications for Higher Education

BEST COPY AVAILABLE

Papers in this series available from NCRIPTAL

Approaches to Research on the Improvement of Postsecondary Teaching and Learning: A Working Paper
Patricia J. Green and Joan S. Stark

Focusing on Student Academic Outcomes: A Working Paper
Joanne M. Alexander and Joan S. Stark

Postsecondary Teaching and Learning Issues in Search of Researchers: A Working Paper
Carol D. Vogel and Joan S. Stark

Classroom Assessment Techniques: A Handbook for Faculty
K. Patricia Cross and Thomas A. Angelo

Teaching and Learning in the College Classroom: A Review of the Research Literature
Wilbert J. McKeachie, Paul R. Pintrich, Yi-Guang Lin, and David A. F. Smith

Psychological Models of the Impact of College on Students
Harold A. Korn

Designing the Learning Plan: A Review of Research and Theory Related to College Curricula
Joan S. Stark and Malcolm A. Lowther, with assistance from Sally Smith

Reflections on Course Planning: Faculty and Students Consider Influences and Goals
Joan S. Stark, Malcolm A. Lowther, Michael P. Ryan, Sally Smith Bomotti, Michele Genthon, C. Lynne Haven, and Gretchen Martens

Faculty as a Key Resource: A Review of the Research Literature
Robert T. Blackburn, Janet H. Lawrence, Steven Ross, Virginia Polk Okoloko, Jeffery P. Bieber, Rosalie Meiland, and Terry Street

The Organizational Context for Teaching and Learning: A Review of the Research Literature
Marvin W. Peterson, Kim S. Cameron, Lisa A. Mets, Philip Jones, and Deborah Ettington

Electronic Information: Literacy Skills for a Computer Age
Jerome Johnston

Design in Context: A Conceptual Framework for the Study of Computer Software in Higher Education
Robert B. Kozma and Robert L. Bangert-Drowns

1987 EDUCOM/NCRIPTAL Higher Education Software Awards
Robert B. Kozma, Robert L. Bangert-Drowns, and Jerome Johnston

Performance Appraisal for Faculty

Implications for Higher Education

Robert T. Blackburn
Judith A. Pitney

Grant Number OERI-86-0010

Joan S. Stark, Director
Wilbert J. McKeachie, Associate Director

Suite 2400 School of Education Building
The University of Michigan
Ann Arbor, Michigan 48109-1259

(313) 936-2748

Acknowledgements

This monograph comprises the ideas and efforts of several individuals at the University of Michigan. The importance of the topic and its relationship to the work of NCRIPTAL was first suggested by Virginia Polk Okoloko. Margery Schnell contributed to the research by summarizing the methods of performance assessment currently used by American business and industry. NCRIPTAL researchers Joan Stark, Marvin Peterson, and Kim Cameron, and OERI Program Director Clifford Adelman provided valuable insights, advice, and criticisms on early drafts of the monograph, as did Sarah Freeman. Professor Martin Maehr, University of Illinois, served as our "outside" critic and provided an extensive and valuable commentary. Mary Joscelyn and Elizabeth Olson have overseen the final editing and production of this publication. Finally, Jane Elliott typed the manuscript, including its several revisions. We are grateful to all of these colleagues for their generous assistance.

Robert T. Blackburn
Judith A. Pitney

Technical Report No. 88-D-002.0
©1988 by the Board of Regents of The University of Michigan for NCRIPTAL

The project presented, or reported herein, was performed pursuant to a grant from the Office of Educational Research and Improvement/Department of Education (OERI/ED). However, the opinions expressed herein do not necessarily reflect the position or policy of the OERI/ED and no official endorsement by the OERI/ED should be inferred.

Contents

Preface vii

I. Theories Underlying Performance Appraisal 1
- A. Introduction 1
- B. Motivation 1
- C. Key Variables Affecting Motivation 5
- D. Summary 6

II. Research on the Outcomes of Performance Appraisal 9
- A. Introduction 9
- B. Influence of Person Perception on Appraisal Outcomes 9
- C. Effects of Gender, Race, and Age in Appraisal Decisions 11
- D. Employee Perceptions of Performance Appraisal 12
- E. Other Specific Effects of Performance Appraisal 14
- F. Outcomes of Performance Appraisal in Educational Settings 15
- G. Salary and Developmental Components of Performance Appraisal 17
- H. Ideas for Improving the Process 17
- I. Conclusions of the Literature Search 19

III. How the Findings Fit with College and University Structures, Norms, Values, and Practices 23
- A. Structures 23
- B. Norms and Customs 24
- C. Values and Beliefs 25
- D. Practices 26

IV. A Recommended Performance Appraisal Plan 31

V. The Feasibility of Change 37

Appendix

Performance Appraisal in Business 41

References 43

Bibliography 51

Tables

1. Approaches to Motivation 6
2. Research Findings on Important Elements in Performance Appraisal 20

Figures

1. Maslow's Hierarchy of Needs 2
2. The Process of Inequity 4

Preface

Criticisms of performance appraisal are currently in vogue. Within the past few years, three ASHE-ERIC monographs have addressed one aspect or another of the topic. Mortimer, Bagshaw, and Masland's (1985) *Flexibility in Academic Staffing: Effective Policies and Practices*, Creswell's (1985) *Faculty Research Performance: Lessons from the Sciences and the Social Sciences* and Boyer and Lewis' (1985) *And on the Seventh Day: Faculty Consulting and Supplemental Income* each deal with evaluating faculty behavior with the aim of changing and rewarding or punishing it. As we prepared this review, an advertisement for the *Journal of Personnel Evaluation in Education* came across our desk that promised to have at least one piece in the first issue dealing with student ratings of college faculty. A recent memorandum from the U. S. Department of Education's Office of Educational Research and Improvement announced a forthcoming request for proposal for National Centers to investigate merit pay as a way of improving instruction, clearly one component of performance appraisal—and one we address. The topic, then, has immediate applied consequences that extend beyond our more general purpose.

This literature review was undertaken to serve NCRIPTAL's principal goal of improving student (cognitive) learning. Program D, Faculty as a Key Resource, contributes to that goal by studying how to change faculty behavior.

We know that faculty want to improve their students' learning. We also know that faculty believe that student learning depends highly on their teaching. We do not know, however, the degree to which that belief is true.

It appears that faculty who add certain strategies to their teaching repertoire increase student learning (McKeachie, Pintrich, Lin, & Smith, 1986). Most faculty, however, do not make major changes in their teaching styles unless they are adequately stimulated or provoked. Consequently, we need to learn under what conditions and circumstances faculty can be motivated to change their teaching behaviors. Appraisal of their performance is a necessary first step.

Our starting point is to ask what is known about performance appraisal. What are its positive and negative consequences? How does the manner in which it is conducted relate to future performance? In what ways do the skills of the appraiser affect the outcomes? To what degree can we expect what is accepted about performance appraisal in other settings to be true for faculty and department chairs in colleges and universities? These and other questions directed our investigation.

This literature review begins with a brief exposition of the principal theories that underlie performance appraisal research in Section I. Section II examines the research studies conducted in the field or in the laboratory to test hypotheses drawn from the theories. The research is principally in the psychological and organizational fields with few contributions from postsecondary education. We end Section II with 15 general conclusions, "truths" we believe postsecondary education needs to take into consideration to improve its performance and more closely achieve its goals.

Section III describes the norms, values, structures, and practices that characterize colleges and universities so that the application of the findings can be inferred as to their likelihood of successful outcomes in postsecondary institutions.

We then advance a plan for performance appraisal well suited for colleges and universities in Section IV and discuss the conditions under which it could be successful in Section V.

The list of cited references, an appendix, and a bibliography complete the monograph. The appendix briefly identifies the main appraisal procedures currently in practice in business. The bibliography contains some references we found useful but were tangential to our aims and hence not cited by us.

I. Theories Underlying Performance Appraisal

A. Introduction

The literature and research relative to performance appraisal covers more than a hundred years of scholarly activity and spans several disciplines, including psychology, sociology, operations research, and organizational behavior. The broad topic of performance appraisal can be divided into three major areas: the supporting behavioral theory, the methodology of performance appraisal, and its effects on performance. While a great deal of attention has been paid to the methodology (format, methods, and physical characteristics) of performance appraisal, somewhat less has been given to the underlying theory. Likewise, it appears that the effects of performance appraisal on individuals and their organizations have emerged as an important research topic only in the last 15 or so years.

Motivation and the variables that influence motivation constitute the central behavioral concept supporting performance appraisal. Section I summarizes the most prominent motivation theories. Our review of the impact or outcomes of performance appraisal (Section II), drawn from a literature base ranging from the highly theoretical or hypothetical to concrete, scientific results supported by data, addresses such issues as stress, mistrust, and increments and decrements in performance.

B. Motivation

An understanding of the topic of motivation is critical to understanding why people behave as they do on the job. Research has demonstrated that the old and simple guidelines for "economic man" are insufficient for understanding human behavior at work. At the same time, the literature shows no "master" theory of motivation or agreement about how motivation contributes to or influences performance.

A basic management principle states that an individual's performance is a function of his or her level of ability and motivation. As organizational performance and its measurement have increased in importance, the level of knowledge and research in the area of work motivation has increased accordingly. One of the simplest and most utilitarian definitions of motivation is that which energizes, directs, and sustains behavior (Atkinson, 1964). Motivation as a concept is a pervasive one, for motivation affects and is affected by a multitude of factors. From the manager's perspective, motivation theory should explain (a) what drives behavior, (b) what direction behavior will take, and (c) how to maintain behavior.

Arousal Theories of Motivation

The concept of human needs as a basic principle of action dominated the study of motivation from the mid-1940s until recent years. While several need theories or variants thereof can be identified in the literature, three deserve special consideration.

The first is Maslow's (1954) theory which argues that individuals are primarily "wanting" creatures motivated by a desire to satisfy certain specific types of needs. The basic proposition of his theory is that human motives can be divided into two classes: those aimed at maintenance and those which are growth-directed. Once a certain need or set of needs becomes satisfied, it loses potency as a motivating force. Maslow also argues that the needs individuals pursue are arranged in an hierarchical order and that every need is inextricably related to those above and below it in the hierarchy (see Figure 1).

Due largely to the popularization of Maslow's model by McGregor (1967), this theory became widely discussed and used by both organizational psychologists and managers. Wahba and Bridwell (1973), however, showed that Maslow's need hierarchy theory has received little clear or consistent support from available research findings.

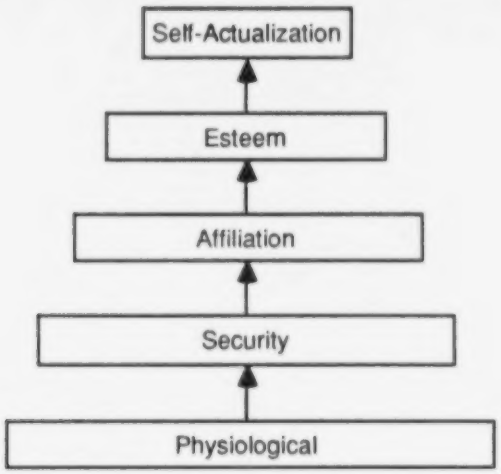


Figure 1. Maslow's hierarchy of needs.

They concluded that Maslow's theory is almost non-testable because of the difficulty in interpreting and putting its concepts into operation.

Alderfer (1969) proposed a modified need hierarchy theory that collapses Maslow's five levels into three—existence, relatedness, and growth needs. He differed from Maslow in suggesting a regression as well as a progression process in the hierarchy. Alderfer contends that when individuals are continually frustrated in attempts to satisfy growth needs, relatedness needs may re-emerge as primary and people may redirect their efforts toward these lower-order needs. He also suggested that more than one need may be operative at the same time.

Before Maslow developed his theory, another important need theory of motivation was developed by Murray in the 1930s. Murray (1938) viewed an individual's personality as being composed of many divergent, and often conflicting, needs that could motivate human behavior. His list of needs includes achievement, affiliation, power, autonomy, nurturance, and deference. The need for achievement has been the subject of a vast amount of research, with the two most prominent contemporary investigators being McClelland and Atkinson. Both have conducted numerous laboratory and field studies that show a strong positive relation between a high need for achievement and high levels of performance and executive success. As they have been developed, McClelland's (McClelland, Atkinson, Clark, & Lowell, 1953) and Atkinson's (1964) achievement models tell us that high achievers:

- like to set their own goals;
- tend to avoid the extremes of difficulty levels in selecting goals; and,
- prefer tasks that provide more or less immediate feedback.

McClelland (1961) pointed out the complex effect of monetary incentives on high achievers. They are unlikely to remain for long in an organization that does not pay them well for exemplary performance. But it is questionable whether an incentive plan actually increases their performance, since they tend to work at peak efficiency anyway. McClelland also noted that achievement motivation does not seem to operate when high achievers are performing tasks that are routine or boring, or where there is no competition. Steers (1983) extensively analyzed Murray's needs theory and the later work of McClelland and Atkinson. He pointed out that situational factors (child-rearing practices, culture, organizational practices) affect the development of achievement motivation in individuals.

In discussing arousal theories of motivation and achievement as a motivator, mention must be made of the related cognitive concept of self-efficacy (Bandura, 1977a, 1980).

Self-efficacy refers to judgments of the likelihood that one can organize and execute given courses of action required to deal with situations that may contain ambiguous, unpredictable, and stressful features. Perceived self-efficacy can have diverse effects on behavior, thought patterns, and affective arousal. Self-efficacy is hypothesized to influence one's choice of activities, effort expended, perseverance when difficulties are encountered, and skillful performance. Bandura asserts that people tend to avoid tasks they believe exceed their coping abilities, but they undertake and perform confidently those activities they judge themselves capable of managing. Personal expectations for success are viewed as important influences on behavior by a variety of theoretical approaches (Bandura, 1982).

A third variant of the arousal theory of motivation we note is Herzberg's (Herzberg, Mausner, & Snyderman, 1959) two-factor theory (developed using engineers and accountants as his test population). It attempts to specify why employees value certain job-related outcomes. Herzberg's is one of the more controversial need theories of motivation because he stresses that some job factors lead to satisfaction, whereas others can only prevent dissatisfaction. Herzberg also states that job satisfaction and dissatisfaction do not exist on a single continuum; the factors producing satisfaction are separate and distinct from those leading to dissatisfaction. Herzberg developed the idea of intrinsic motivators (achievement recognition, the work itself, responsibility, advancement, growth) and extrinsic motivators or "hygienes." Intrinsic factors relate to the content of the job itself; extrinsic hygienes are largely determined by the organization (e.g., salary, co-worker relations, supervisory style). There have been numerous studies of Herzberg's theory, which is frequently cited in the literature. According to Harlan, Kerr, and Kerr (1977), the research designed to test it has not provided clear evidence either supporting or rejecting it.

In recent years there has been a shift away from need-based theories of arousal (Weiner, 1972) to theories suggesting that people are aroused by the presence of others and the knowledge that other people are evaluating them (Mitchell, 1982).

Choice Theories of Motivation

The theories of Maslow, Murray, and Herzberg provide some understanding of the arousal and energizing aspects of the motivational process, but they do not explain why people choose a particular behavior to accomplish work-related goals. Four major "choice" theories of motivation have been developed in the last 30 to 40 years to explain why people start, direct, sustain, and stop behavior. Whereas "arousal" theories explore the needs that motivate behavior, "choice" theories discuss how needs and other factors, such as values and perception of the situation, interact and produce certain kinds of behavior.

The first of these, expectancy theory, was formulated by Vroom (1964) and is the most popular cognitive explanation of worker motivation. It states that motivation is a combined function of (a) the individual's perception that effort will lead to performance and (b) the perceived desirability of outcomes that may result from performance. The expectancy theory model has three key elements:

1. Expectancy—the belief that a particular level of effort will be followed by a particular level of performance.
2. Valence—the measure of an individual's feelings about a particular outcome.
3. Instrumentality—the relationship between performance and outcomes.

According to expectancy theory, the force on an individual to engage in a particular action is a multiplicative function of these three elements. The theory tells us is that the more likely people feel that a level of performance will lead to desired outcomes, the more likely they will be to try to perform at that level.

Porter and Lawler (1968) have extended the principles of the basic expectancy model to examine the factors that influence an employee's job performance and satisfaction.

They argue that satisfaction is an effect rather than a cause of performance. Different levels of performance determine rewards that, in turn, produce variations in employee satisfaction. Considerable research on the implications of the expectancy theory for managers has been conducted. Nadler and Lawler (1977) recommend that managers identify what outcomes each employee values. They also urge managers to link the outcomes desired by employees to specific performance desired by the managers. Nadler and Lawler tell managers to make sure changes in outcomes or rewards are sufficiently large to motivate significant behavior. Trivial rewards result in minimal amounts of effort and thus only slight performance change.

Equity theory (Carrell & Dittrich, 1978; Goodman, 1977) is a theory of social comparison processes which suggests that people are motivated by a desire for fairness. Equity theory is based on the relationship between inputs and outcomes. Steers (1981) outlines the motivational process of inequity in Figure 2.

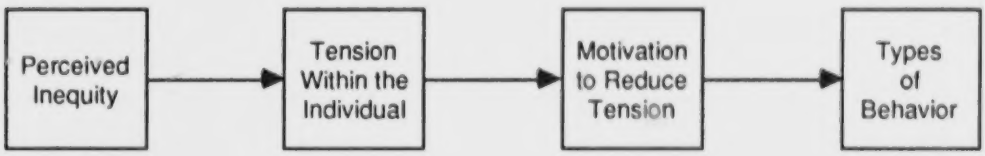


Figure 2. The process of inequity.

When individuals believe they have been treated unfairly (e.g., not given a sufficient monetary reward), they change inputs or outcomes to restore equity. Although over-reward and under-reward are similar from a theoretical perspective, research suggests otherwise. People are more comfortable with over-reward than with under-reward.

A third important choice theory of motivation is the goal-setting model. Locke (1968) deals with the interplay between an individual's conscious or unconscious goals and task performance. In goal-setting theory the cognitive determinants of behavior are values and intentions. Researchers have found that goal setting not only affects performance levels (i.e., people with goals work harder than those without goals), but it also directly affects satisfaction. Goals that incorporate specific performance standards lead to higher performance than general goals (Locke, 1978). Proximal goals, which are close at hand and can be achieved quickly, result in greater motivation and higher performance than goals extending far into the future (Bandura, 1977b). Current research on goal setting theory looks at whether participative or assigned goal setting works better and whether rewards directly influence motivation and the level of the goal.

A fourth theory of motivated choice is the reinforcement theory, which is founded on the idea that voluntary human behavior is environmentally determined. The consequences of a given behavioral act determine the likelihood that a particular behavior will be engaged in again. Rewarding consequences inform and motivate. Skinner (1971) is perhaps the best known proponent of the reinforcement, or operant conditioning, paradigm. While many researchers have expressed concerns about the ethical considerations of using reinforcement principles, these principles (e.g., behavior modification, positive reinforcement) are widely used in industry.

Some data, however, call into question the reward principle. For example, Atkinson's theory of achievement motivation and the empirical findings on the consequences of success and failure demonstrate that people seek moderately difficult tasks. Hence success at an easy task should not predict the probability of undertaking that task again. There is also a growing literature on whether the use of extrinsic rewards may lower intrinsic motivation on a positively perceived task.

Intrinsic motivation, as defined by Deci and Ryan (1985), refers to the motivation to perform a task or activity when the reward is satisfaction derived from performance. Extrinsic motivation, on the other hand, is defined as the motivation to perform a task or activity that derives strictly from the external inducements or incentives provided by others.

Recent research has sought to determine whether external rewards act on intrinsic motivation in an additive fashion. Some evidence suggests that external rewards reduce intrinsic motivation (e.g., Deci, 1971, 1972, 1976; Pinder, 1976). Other researchers differ in their conclusions. Calder and Staw (1975) demonstrated that although monetary rewards tend to decrease intrinsic motivation on interesting tasks, rewards may actually increase intrinsic motivation on boring tasks. Deci (1976) suggests that if financial rewards convey that the individual is being paid for competence, intrinsic motivation may not be decreased and could even be increased by such rewards. Boal and Cummings (1981) tested Deci's hypotheses in a natural work environment. Their results lend some support to his argument that performance-contingent reward systems may decrease intrinsically motivated behavior but they do not support Deci's explanatory frameworks.

C. Key Variables Affecting Motivation

Porter and Miles (1974) identify four major categories of variables that influence motivation: (a) individual characteristics, (b) job characteristics, (c) work environment characteristics, and (d) the characteristics of the external environment.

Several individual characteristics can significantly influence a person's motivation and performance. As pointed out previously in this section, there is fairly consistent evidence that individuals who have higher needs for achievement generally perform better than those who have lower needs. Individual needs for social and economic status are an important factor influencing motivation. Likewise, one's interests and attitudes also play a role in motivation and performance. Korman (1971) maintains that individuals attempt to behave in a fashion consistent with their own self-image.

The second major set of variables affecting motivation involves what the person does at work—the nature of the job. The variety of activities required to do a job, the significance and challenge of the tasks, the manner in which the supervisor organizes and defines worker activities, the type of intrinsic rewards, and the amount and type of feedback one receives as a consequence of performing the job all influence motivation. Several studies (e.g., Hackman & Oldham, 1976; Steers & Porter, 1974) have found that enriching an employee's job by allowing more variety, autonomy, and responsibility can result in improved performance. But this is not true for everyone—some individuals do not want more responsibility or challenge.

The nature of the work environment is also relevant to the motivational process. The nature of peer-group interactions can significantly influence an individual's effort; supervisory or leadership style influences motivation; and the reward practices of the organization and its openness of communication can affect motivation. Another aspect of the environment that affects the comfort and motivation of professionals is organizational climate. Argyris and Schön (1974) distinguish between Model I (decisions made at the top; little feedback) and Model II (open communication and feedback encouraged) systems.

The fourth major category of variables that can affect employee motivation is the external environment. According to Perry and Porter (1982) this category can be subdivided into socio-normative, political, demographic, economic, and technological concerns. For example, to the extent that the general public holds unfavorable attitudes about colleges and universities and their faculties, motivation-relevant perceptions, such as self-worth and personal significance, can be affected.

In addition to the Porter and Miles framework of factors that influence the motivational process, one needs to consider the relationship between aging and motivation. Few studies have been specifically concerned with motivational changes in adulthood (Maehr & Braskamp, 1986). Research has indicated that adult motivation relates to four factors: psychological career, time orientation, sources of value, and motivational tendency. Raynor and Entin (1982) analyzed the effects of success and failure on motivation for each new subsequent activity along a career path. They point out that the distinction between open and closed career paths is critical to understanding adult personality functioning. The closed career path has a final or ultimate goal that, when attained, will mark the end of striving along the path. This is because the last goal of the career is fixed at the outset and remains unchanged as a function of success in moving toward it. On the other hand, an open path may initially have an ultimate goal, but an immediate success suggests new goals that add on to the end of the path, whose length now remains the same or is even increased.

Open and closed career paths are not fixed for a given individual since an open path can become closed and a closed path can become open. But the older person typically has a constricted future path. Positively motivated people who do have future career paths in the later stages of life are, by this theory, expected to experience fewer motivational problems than their "futureless" contemporaries, because moving toward future paths is something they have done all their life.

While more research dealing with the questions of age and motivation is needed, it seems reasonable to assume that advancing age does not necessarily mean less desire either for challenge and achievement or for rewards and recognition.

D. Summary

Table 1 outlines the seven major theories of motivation discussed in Section I. As has been indicated, the literature on these and other theories of motivation is voluminous. Steers and Porter (1983), Mitchell (1982), and Landy and Becker (1987) provide comprehensive reviews of the major theories and the current research associated with them. So do Weiner (1972) and Landy, Zedeck, and Cleveland (1983). Maehr and Braskamp (1986) advance the concept of personal investment as central to the understanding of motivation. Staw (1984) cautions that there are many limitations to motivation theories in both the conceptual and applied senses. He notes that applying such theories, which are derived primarily from experience in business and industry, to the university may be inappropriate, a point we address in Sections IV and V.

There seems to be agreement that the emphasis of future research on motivation should be on the additive or interactive effects of the various approaches and on the thought processes that are the immediate antecedent of motivation. Other important

TABLE 1

Approaches to Motivation

APPROACH	CHARACTERISTICS	THEORIES	HOW IT WORKS
Arousal	Concerned with factors that start or arouse motivated behavior.	1. Need Hierarchy (Maslow, Alderfer) 2. Achievement Motivation (McClelland, Atkinson) 3. Two-Factor (Herzberg)	Motivation is started and sustained by satisfying individual needs (money, status, achievement, working conditions).
Choice	Concerned not only with things that start behavior but also with the choice of behaviors and factors that increase the likelihood that desired behavior will be repeated.	1. Expectancy (Vroom) 2. Reinforcement/ Operant Conditioning (Skinner) 3. Equity (Goodman) 4. Goal-setting (Locke)	Motivation is started and sustained by clarifying the individual's perception of work inputs and by rewarding desired behavior.

issues to be addressed include which motivational approach works best in which situation(s) and how group processes affect motivation.

It is important to recognize that the beliefs individuals hold about motivation can have important implications for their attitudes and behaviors. In the case of faculty evaluation, a department chair's or dean's view of motivation could well mediate the process. Equally important is the fact that personal characteristics unique to individuals can have a significant impact on their work behavior. There is no universally effective way of motivating individuals. The importance of recognizing individual differences among employees (faculty) cannot be over-emphasized. Most managers recognize variations in employees' needs, abilities, and traits. They also must be aware that different employees have different preferences for the rewards available for good performance. Somehow department chairs and deans must fit the work situation to the person in order to maximize faculty effort.

II. Research on the Outcomes of Performance Appraisal

A. Introduction

As noted previously, the literature on the outcomes of performance appraisal has burgeoned in recent years. While some of it is supported by field or laboratory studies, a considerable amount is based on untested hypotheses. This review focuses on the former category and addresses (a) what is known about factors that influence the process of performance appraisal; (b) the effects of the process on those individuals and organizations participating in it; and (c) what ideas have been advanced for improving the process.

B. Influence of Person Perception on Appraisal Outcomes

Recent research on the processes involved in making a judgment about performance has concentrated on identifying the information acquisition, information integration, and judgmental process that raters use in making an appraisal. Several distinct literatures address issues of perceiving and judging the behaviors of others. These have been labeled (a) attribution theory, (b) implicit personality theory, and (c) social cognition (Ilgen & Favero, 1985). All deal with some aspects of the broader topic of person perception, which is the basis of performance appraisal.

The social psychological research on person perception concludes that perceivers assign persons to categories defined by resemblances among their members and exemplified by category prototypes or images. People may be assigned to categories automatically by virtue of their possession of obvious attributes (e.g., gender or race). When no category provides a satisfactory fit or when there is a discrepancy between available information and the initial categorization, automatic categorization may be superseded by a controlled process described by attribution theory.

Attribution theory is based on the idea that an observer interprets behavior in terms of its perceived causes—that is, the observer attributes a behavior to some cause. Attributions refer to the explanations people develop about why people behave as they do or why particular events occur. Attribution theorists postulate that, in achievement-related contexts, success or failure is cognitively attributed to such factors as ability, effort, task difficulty, or luck (Weiner, 1972).

In summary, to the degree the behavior of an employee is consistent with the supervisor's expectations, it is noted and stored automatically. It is only when a behavior departs from expectations, or when the task is somehow changed, that conscious attention and recognition (attribution) processes are engaged in by the supervisor.

When an employee is assigned to a category, further memory-based judgments of that employee are colored by the category prototype. This process, which is functionally identical to stereotyping, can produce either under-evaluations or over-evaluations of employees by associating the general evaluation of the category with the person, producing false memories of the person, or both (Feldman, 1981). Dipboye (1985) has described the stereotype-fit model in which raters possess stereotypes of the ideal occupant of a job and their evaluations reflect their perceptions of the goodness of the fit of the employee to the job. Feldman (1981) argues that categorization affects performance appraisal by limiting and selecting information about the employee when memory-based judgments are made and by influencing stimulus-based judgments through the operation of attributional bias.

Implicit personality theory (IPT) concerns persons' perceptions of the relationship between traits, or how traits co-vary in others (Schneider, 1973). Two performance appraisal rating errors based on IPT are systematic distortion and halo. The term halo "implies inflation of correlations among dimensions, while systematic distortion

implies that raters make memory-based ratings, with correlations among dimensions biased in the direction of semantic similarity" (Borman, 1983, p. 132).

Researchers in the field of performance appraisal have devoted much attention to the problems of inadvertent judgmental biases that seem to be an inherent feature of rating processes and outcomes. In the view of Nisbett and Wilson (1977), halo error is the longest recognized, most pervasive, and least understood form of rating error.

The major contribution of the research on IPT is its emphasis on accuracy as a dependable variable. The work of Borman (1979) and of Bernardin and Pence (1980) clearly shows that reductions of halo and stringency or leniency errors do not necessarily indicate greater accuracy, that is, ratings with fewer overall errors.

Social cognition research looks at the cognitive processes involved in person perception. It views the perceiver (rater) as an information processor and is concerned with how information is acquired, stored, and retrieved. A number of biases may affect the processing of performance appraisal information at the early stages of information processing. One line of social cognition research on performance appraisal looks at the nature of information stored relative to the purpose to which the information is to be put. Hamilton, Katz, and Leirer (1980) suggest that the rater's beliefs about the use to which the performance information will be applied (e.g., counseling or administrative purposes) affect the nature of the information that is collected, combined, and recalled. Consistent with this, Zedeck and Cascio (1982) demonstrated that the purpose of performance appraisal affects the way raters combine information. Williams, De Nisi, Blencoe, and Cafferty (1985) also addressed this important issue and concluded that evaluators search differentially for more comparative information when they have to select one of several subjects for some treatment.

In addition to appraisal purpose, many other factors may influence raters' cognitive activities and their abilities to assign accurate performance appraisals. Much theoretical and empirical research suggests that a first or early (e.g., first six months on the job) impression may bias the recording of performance-related behaviors. Several researchers (e.g., Balzer, 1986) have argued that initial impression leads to a confirmatory observation bias for behaviors consistent with the initial impression.

In a longitudinal study, Hogan (1987) explored the question of whether supervisors' expectations about subordinates' performance might bias their subsequent evaluations of those subordinates. Hogan's respondents were supervisors in a large West Coast bank. Results indicated that when a subordinate's actual performance disappoints a rater's expectations about that performance, subsequent ratings will be lower than actual performance warrants. When actual performance exceeds a rater's expectations about that performance, subsequent ratings will again be lower than the actual performance warrants. Hogan concludes that raters are averse to being wrong and punish the source of their disconfirmation.

Ilgen and Favero (1985), in reviewing what social psychology research on person perception might tell us about performance appraisal, caution that the literature to date has not paid enough attention to three critical features of performance-appraisal: (a) future interactions between the rater and ratee; (b) rater/ratee interdependence; and (c) ratee behaviors versus the consequences of those behaviors.

In relation to points one and two, some research shows that decision makers who commit themselves to a particular course of action may make subsequent related decisions in a non-optimal manner to justify the previous commitment (Staw, 1976, 1981). Bazerman, Beekun, and Schoorman (1982) discuss the relevance of Staw's findings to performance appraisal. They suggest that a rater's subsequent decisions about someone he or she previously decided to promote may be biased systematically in one or more ways. First, the raters' perception of information may be biased by their previous decision—that is, the raters pay more attention to information that supports their promotion decision than to information that discredits it. Second, even if raters privately judge the ratee's performance negatively, they may not give a negative

evaluation because it would publicly contradict the previous commitment to the ratee. Unfavorable evaluations expose not only the subordinate but also the supervisor to criticism.

Other factors that might moderate the accuracy of performance appraisal include the rater's style of management and his or her perception of how the work should be done.

C. Effects of Gender, Race, and Age in Appraisal Decisions

There is a modest body of research addressing how gender, race, and age are factored into the performance appraisal process and how they cause systematic biases. From the preceding discussion of person perception, it is clear that once a person is categorized, recall and recognition of that person are biased toward the general characteristics of the category. Taylor and Fiske (1978) have shown that aspects of the situation that make a given feature (e.g., race, gender, age) of a person more salient in the perceptual sense strongly influence categorization (e.g., one woman in a work group of twenty).

Schmitt and Lappin (1980) tested the hypothesis that people rate those similar to themselves with more confidence, which is reflected in larger variations in performance ratings. Their study, which used undergraduate psychology students as the subjects, concluded that people are more confident rating others in their own racial group than they are of ratings of other racial groups. However, their data did not support a similar conclusion for gender subgroups. Other experiments examining race effects have yielded inconsistent findings: some have shown bias against blacks (Hamner, Kim, Baird, & Bigoness, 1974) and others show no differences in ratings as a function of race (Maruyama & Miller, 1980).

The effects of gender on effort attributions in performance evaluation have been considered by a number of researchers. Deaux and Emswiller (1974) concluded that in masculine gender-typed tasks, identical performance by males and females is perceived to be caused by different factors. Male performance is attributed more to individual characteristics such as ability than is comparable female performance. Female performance is attributed more to extra-individual factors such as luck. These differences in attributions may disadvantage the females in performance evaluations where personal characteristics, such as effort, are weighed heavily and where the evaluator's knowledge of the subject's job performance is limited.

In looking at managerial performance evaluation, Rose (1978) examined the effects of three gender variables—sex of the rater, sex of the subject, and sex of the subject's subordinates. His research population was 86 graduate students in business. Rose found that, despite comparable performance evidence, both male and female subjects attributed greater effort to managers whose subordinates were predominantly of the opposite sex than to managers whose subordinates were predominantly of their same sex. Rose's results extend previous attribution research by demonstrating that gender effects on attributions depend not only on the subject's sex, but also on the subject's sex interacting with the larger sexual context of the job.

Wexley and Pulakos (1982) examined the effects of rater gender and/or ratee gender on performance ratings in four kinds of field settings that encompassed 17 different organizations. In contrast to the research on race bias, they showed that female employees receive similar ratings. Male managers and subordinates did not produce more variable performance ratings when appraising other males than when rating females. Wexley and Pulakos' results involving female managers and subordinates produced ratings that were significantly less variable when appraising other females than when appraising males. They speculate that the rating consistency is due to the fact that the women may be uncertain about role expectations for women managers and therefore restrict the spread in their ratings by consistently using the middle of the rating scale.

Other studies on bias against women have found that women are rated as highly as, or higher than, men when both sexes exhibit high levels of performance, and less favorably than men when both sexes exhibit mediocre or poor levels of performance (Abramson, Goldberg, Greenberg, & Abramson, 1978).

The effects of appearance on how individuals and their work are regarded and how rewards are allocated in work settings has also been examined in the psychological literature. Heilman and Stopeck (1985), in a laboratory experiment, demonstrated that attractiveness had both a favorable and a deleterious effect for working women. In performance evaluations, attractiveness generally benefited females in non-managerial positions but always proved to be a liability for female managers. Males did not prove to be vulnerable to bias on the basis of appearance. Heilman and Stopeck's findings are consistent with the idea that differential assessments of person-job fit are a key element in the performance appraisal process.

The amount of published research on age and performance appraisal is very small. Using 513 managers in a large manufacturing company as their subjects, Cleveland and Landy (1981) found that older employees received lower ratings than younger employees on the specific appraisal items of self-development and interpersonal skills. Rosen, Jerdee, and Lunn (1981) reached the same conclusion. However, both studies conclude that the age of the rater or ratee did not seem to exert a systematic distorting influence on overall performance ratings.

Schwab and Heneman (1978) point out that if a rater holds stereotypes of the ratee based on the ratee's age, then an inaccurate performance assessment will likely be made.

D. Employee Perceptions of Performance Appraisal

Turning to the effects of the appraisal process on those participating in it, several theorists have written on the potential negative effects of performance appraisals on employees' subsequent attitudes and behaviors. McGregor (1957) argued that performance appraisal creates a situation in which the supervisor must behave in a threatening, authoritarian, and ego-deflating manner with those employees judged to be poor performers. Others (e.g., Patz, 1975) have described the conflict and misunderstanding that arise during performance appraisal. Taylor, Fisher, and Ilgen (1984) state that even when an appraisal system is perceived as fair, negative feedback may threaten employees' perceived freedom of choice and can result in defiant opposition or reaction to the supervisor. Similarly, Meyer (1975) suggested that employees given "below average" merit ratings become alienated and demoralized. However, the research supporting these kind of conjectures has been scant until recently.

Pearce and Porter (1986) report on the results of a large-scale research effort on the attitudinal impacts of formal performance appraisal feedback. Their research data were collected from separate samples of managers and non-management employees in two federal agencies over a 30-month period. They concluded that performance feedback indicating that a person is "satisfactory" or "meeting standards" is experienced as negative by many recipients, not just by those receiving objectively poor performance ratings. In Pearce and Porter's sample, this outcome was true for the managers who were on a merit pay system but not for the non-management employees whose pay was not affected by their rating. Another important finding of this study was that the receipt of relatively low ratings caused a distinct and significant drop in attitudes toward the organization within two months of feedback of the appraisal results.

Pearce and Porter caution that any appraisal system that provides data, implicit or explicit, on how one ranks compared to one's peers is likely to be dysfunctional and to generate some loss in positive feelings on the part of those who are not in the upper part of the distribution. This suggests that organizations need to consider carefully

how appraisal systems, even when technically correct, may affect the attitudes and performance of their "solid citizens" who are performing at acceptable, but not outstanding levels.

In a sample of classified university employees, Dorfman, Stephan, and Loveland (1986) explored three dimensions of performance appraisal: two developmental dimensions (being supportive, emphasizing performance improvement) and one administrative dimension (discussing pay and advancement). Results suggest that supervisors exert more efforts to clear up performance problems with low-performing rather than with high-performing subordinates. Dorfman et al. found that the developmental factor of supervisor support (e.g., taking the attitude of a helper) was positively related to employees' motivation to do well in the future. They also found that discussing pay and advancement promoted positive feelings toward the appraisal process. However, and perhaps most significantly, Dorfman et al. found that performance appraisal behaviors did not have a positive impact on subsequent job performance. In this their findings are consistent with several reviews of appraisal research that have found few examples of constructive actions or significant performance improvements resulting from appraisal interviews (Bernardin & Beatty, 1984).

Landy, Barnes, and Murphy (1978) and Landy, Barnes-Farrell, and Cleveland (1980) researched the question of whether perceptions of the fairness of an appraisal system are affected by the level of previous performance rating; that is, are employees who receive high ratings more likely to describe the evaluation process as fair and accurate than those who receive low ratings? Their conclusion, based on a sample of almost 300 mid-level managers in a manufacturing company, was no. In their view, perceptions of fairness of performance appraisal are a function of the components of the evaluation process itself—especially its frequency, the supervisor's knowledge of the performance and duties of the subordinate, the subject's opportunity to express feelings when evaluated, and the identification of action plans in relation to performance weaknesses.

Fulk, Brief, and Barr (1985) provided a cross-validation and replication of the Landy et al. research. Their sample was 198 engineers in the telecommunications division of a large electronics firm. Fulk et al. concluded that perceptions of fairness and accuracy in performance appraisal may depend as heavily on the level of trust in the ongoing superior-subordinate relationship as on the characteristics of the performance appraisal process itself.

On the other hand, Stone and Stone (1984) concluded that the favorability of feedback does affect the perceived accuracy of such feedback. Their findings, based on a laboratory study using 80 undergraduates as subjects, agree with those of Shrauger and Schoeneman (1979). If their results are correct, it suggests that when poor performers are given accurate feedback about their performance, such feedback may be perceived as inaccurate. Thus it may be extremely difficult to alter the behavior of poor performers through the use of appraisal feedback.

In the study of a research-and-development organization, Dipboye and de Pontbriand (1981) found that perceived favorability of the appraisal, opportunity to state one's own side of the issues, being evaluated on relevant job factors, and discussing plans and objectives related positively to opinions of the appraisal and the appraisal system. Although negative feelings due to negative evaluations may not be eliminated entirely, actions on the part of the supervisor to enhance perceptions of the appraisal process may increase employee acceptance of feedback. Dipboye and de Pontbriand also concluded that employees' opinions of the appraisal system may be as important to the long-term effectiveness of the system as is the validity and reliability of the system. Shields' (1984) report on a Canadian public sector firm employing 60,000 individuals confirmed that manipulation of the performance appraisal variables cited by Dipboye and de Pontbriand did constitute a reliable formula for designing a performance appraisal system that is trusted and accepted by employees.

Mount (1983) assessed managers' and employees' satisfaction with a performance appraisal system in a large multinational corporation. His data show that perceptions of the appraisal system are moderated to a certain extent by the role of the individual in the organization. Employees' satisfaction with the appraisal system was accounted for by their overall experience with it, the quality of the appraisal discussion, the way in which the rating forms helped to discuss performance, and the way the forms helped formulate development plans. For the managers, however, their satisfaction with the process focused on how well it helped communicate to the employees how they were performing.

Ilgen, Peterson, Martin, and Boeschen (1981) explored the process of supervisor-subordinate feedback sessions for exempt employees of a wood products industry headquartered in the Northwest. Their results show that, even when the feedback was straightforward and presented on a scale with which employees were very familiar, employees still overestimated their own performance. Explanations given to subordinates about their performance are a likely contributor to this problem. Trying to avoid the unpleasantness of negative feedback, supervisors often may present a given performance rating and then proceed to explain away the low rating to improve the interpersonal climate of the appraisal session. Ilgen et al.'s research project also demonstrated that reactions to appraisal sessions do not occur in a vacuum. They fit into the total perception of the work experience. This study indicated that employees are likely to respond better to performance feedback when they receive it in a consistent and considerate manner throughout the year as opposed to only once a year.

The question of how the performance appraisal process influences commitment and climate within an institution requires research beyond that of Pearce and Porter (1986). The studies that focus on the antecedents of employee commitment to organizations suggest that commitment is influenced by four sets of factors, one of which is work experiences (Morris & Steers, 1980; Mowday, Porter, & Steers, 1982). It seems reasonable to conclude that the results of a performance appraisal system—part of the work experience—have the potential to influence commitment to some degree. If the appraisal system is successful in increasing employee motivation, this too may influence commitment by activating an employee's achievement motive (Steers, 1977). Steers and Lee (1983) argue that performance appraisal outcomes also have a reciprocal relationship with organizational climate: the extent to which the appraisal system is considered fair and equitable bears on professional esprit, the level of conflict, and the stature of those in leadership positions.

E. Other Specific Effects of Performance Appraisal

The consideration of the effects of performance appraisal systems on employees' attitudes and commitment has been extended to include such issues as trust, stress, anxiety, and turnover. For the most part, the literature here addresses untested hypotheses.

Cummings (1983) speculates that well-developed appraisal systems will clarify the basis for an individual's judgments on the degree of trust he or she should have in a supervisor or in the organization.

According to Ivancevich and Matteson (1983), the nature of stress and its impact on performance in organizational settings involves a complex set of issues that few researchers have considered. They argue that a moderate amount of stress and a moderate amount of task difficulty have positive performance effects.

There seems to be consensus that stress build-up occurs when motivating conditions on the job are not improved by the person's actions (French, Caplan, & Harrison, 1982). The performance appraisal process is only one of many factors that may increase or decrease stress, and it is one of which managers should be aware. The stress-related outcomes of performance appraisal will differ by employee: the task

relevant self-images of employees who see themselves succeed on their jobs improve, while the self-images of those who do not do well worsen. This lowered self-image may lead to increased tension and anxiety that, when they go beyond a certain optimum point, may cause a further degradation in performance.

Most studies of job turnover have examined its relationship to job satisfaction and have reported a consistent negative relationship (Locke, 1976). It is highly probable that other variables mediate the relationship between job satisfaction and the act of quitting. One of these could be the performance appraisal process. It is also important to note that employees may respond to job dissatisfaction with less extreme measures than quitting (e.g., engage in other kinds of withdrawal behavior such as absenteeism, passive job behavior, reduced level of involvement). The relationship of the outcomes of performance appraisal to job turnover is at this time a researchable question. While managers usually want performance appraisal to bring about a turnover of poor or low-performing employees, they need to know if the particular system they follow is also causing employees with good, acceptable levels of performance to leave.

F. Outcomes of Performance Appraisal in Educational Settings

In contrast to the numerous volumes on how to evaluate faculty ("cookbooks"), the literature on the outcomes of performance appraisal in educational institutions is sparse. Some observers (e.g., Raelin, 1985) have commented on the built-in conflict between the nature of professionals, such as faculty, and organizational control systems, such as performance appraisal. Based in part on a claim for autonomy in their work accompanied by highly developed and specialized skills, professionals expect to direct their own activities in the workplace free from interference by others. While professionals may submit to evaluation of their performance in the process sense, they normally believe that only their peers are qualified to evaluate their work.

Barber (1983), in a report for the Education Commission of the States, surveyed teacher evaluation and merit pay plans looking for successful models. He points out that teacher evaluation systems seem to suffer from a confusion of purposes; they are split between those that only reward or punish teachers (summative) and those designed to improve teaching performance (formative). But unified evaluation systems are rare and, when they work well, they usually involve peer-mediated self-appraisal as the central element. Barber reports that attempting to improve K-12 teacher performance through a reward-punishment system alone is nearly always unsuccessful.

Miller and Young's (1979) survey of the literature on the effects of merit pay policies in educational institutions concludes that the limited data available is not encouraging. Like Barber, they chronicle a history of numerous, mostly unsuccessful attempts to make merit pay work in the public schools.

Both Barber's and Miller and Young's conclusions on the effects of merit pay plans on teaching are consistent with findings in other areas. For example, Pearce, Stevenson, and Perry's (1985) study of the Social Security Administration found that its merit pay program for managers had no effect on organizational performance. While recognizing that the implementation of merit pay systems may be dysfunctional, many scholars continue to advocate performance-contingent pay. Pearce et al. and Barber disagree and argue that the whole issue needs re-examination.

While not a performance appraisal study, Hind, Dornbusch, and Scott (1974) make a distinct contribution by applying a theory of evaluation and authority to the system used for evaluating faculty at Stanford. Interviews with a random sample of 100 professors show that satisfaction with evaluation increases with perceived agreement among evaluators. Agreement, in turn, is a function of a field of knowledge's dependence upon a body of theory.

Watts (1985) analyzed the impact of a performance-contingent reward system on faculty research and publication activity in the management department in a school of business at a Southwestern university. His results indicate the ineffectiveness of the system, and he concludes it did not work for two main reasons: (a) when goals were set, not enough attention was paid to previous performance levels and this led to under-utilization of resources; (b) the university reward system, which was based on extrinsic rewards, did not optimize faculty motivation. Watts argues that if administrators want faculty to be more prolific researchers, they should examine the rewards that are intrinsic to research activity (e.g., challenge, expertness, discovery) and create situations in which these rewards can be obtained.

Terpstra, Olson, and Lockeman (1982) investigated the impact of an MBO (management by objective) application on 23 faculty in the College of Business and Economics at the University of Idaho with questionnaire data on performance perceptions and satisfaction prior to and one year after the introduction of the plan. There were additional data on research output. The results suggest that performance increased with MBO while satisfaction generally declined.

Ehli (1986) examined faculty attitudes toward the merit bonus plan administered in South Dakota's six public colleges and universities for the 1983-84 academic year. His survey population was 25 percent of the faculty in each of the six institutions. Ehli reports that less than 40 percent of the respondents favored the merit concept and more than 50 percent indicated that the merit bonus plan had a negative effect on faculty cooperation. Nearly 62 percent believed the plan had a negative effect on cooperation between faculty and the merit decision makers.

Holley, Halpin, and Johnson (1982) examined faculty views of the performance appraisal process in a large Southeastern university. Their questionnaire, administered to 60 department heads and 900 faculty, covered topics concerning characteristics of the appraisal program, its perceived effectiveness, and recommended changes. Holley et al. found that the department heads' assessments of how well they administered the performance appraisal program were significantly more favorable than the faculties' assessments. The divergence of opinion between department heads and faculty was reflected on 12 of the 15 items pertaining to how the appraisal process was conducted. Faculty were less certain than department heads that rewards were commensurate with performance, that serious efforts were made to reward outstanding performance, and that merit increases were given in accord with actual performance.

Ormrod (1986) reports a similar set of reactions by the faculty at the University of Northern Colorado when that institution implemented a new faculty evaluation plan in 1983-1984. Of the 435 faculty members evaluated, 95 (22 percent) appealed their evaluation. Inconsistency among evaluators (department chair, dean, vice president) was a significant factor in faculty dissatisfaction.

D'Heilly's (1975) dissertation compared faculties of two small private junior colleges, one of which introduced a comprehensive evaluation system. At the end of a year, faculty at the experimental institution believed that a negative change had occurred in supervisor/supervisee interactions and in their attitudes toward evaluation of their performance, whereas there was no change at the control college.

In a supporting study, although a simulation rather than an actual experiment, Hobson, Mendel, and Gibson (1981) found that, even when criteria were established by both faculty and the chair, when 20 faculty in psychology rated a series of hypothetical performance profiles, the criteria they weighted most heavily differed from what the chair did. Furthermore, the chair was not aware of what behaviors he was giving the most importance.

While focusing on evaluation rather than performance appraisal, two other studies merit inclusion. Ory and Braskamp (1981) had University of Illinois faculty respond to three (simulated) student evaluation reports on their potential for accuracy,

trustworthiness, usefulness, comprehensiveness, believability, interpretability, and value as information for self-improvement and promotion purposes. Faculty said the information was more credible, useful, and accurate for their self-improvement than for promotion purposes.

Kasten (1984) also used simulated faculty profiles of teaching, research, and service so that each role had a strong, intermediate, and weak evaluation resulting in nine different combinations. A sample of 135 tenured faculty in social science departments in a Midwestern research university gave judgments on the promotability and percent merit raise for each of the nine cases. While research was the most rewarded role, other activities came into play, as did career stage (near retirement) and administrative service, both of which led to higher raises. The size of the department also mattered inasmuch as smaller units tended to be more egalitarian.

G. Salary and Developmental Components of Performance Appraisal

One issue vigorously debated in the literature is whether or not the administrative and developmental functions of performance appraisal should be separated to minimize the conflict between the helping and judgmental roles of the supervisor. Ilgen and Feldman (1983) suggest that it is naive to think that employees could ever separate the functions, even if it was possible organizationally to do so. Similarly, Prince and Lawler (1986) and Dorfman et al. (1986) found no detrimental or interactive effects between salary discussions and developmental processes during the performance appraisal. Their research populations were, respectively, a large multi-industry company and a university. While in these two cases the conclusions are research based, much of the argument on the other side (namely, advocating the split) is theory based. The theoretical argument supporting the split role has to do with role incongruity. How can the supervisor be a helper and a judge at the same time? It is suggested that discussing salary can cause the subordinate to be defensive, closed, and mistrustful.

Xerox's Reprographic Business Group (RBG) seems to have agreed with the split-role advocates. Deets and Tyler (1986) report that performance coaching, when attempted in an environment that included discussion of merit increases, was often unsuccessful. RBG's new performance appraisal system has the merit increase discussion take place approximately one to two months following the appraisal discussion.

Finn and Fontaine (1984) conducted a study of the human services department of a state government. They did so at the request of top management who knew that their managerial and professional personnel were especially dissatisfied with the system of performance appraisal in place. Finn and Fontaine interviewed 111 professionals in individual sessions lasting from one to one-and-a-half hours. The majority of the people interviewed wanted their goal-setting process to take place at a time different from the appraisal. They felt this change would reduce their negative feelings toward performance appraisal.

It seems clear that future research needs to test the compatibility/non-compatibility issue more extensively. For those individuals who respond to goal setting and monetary incentives, combining the discussion of performance and pay makes sense. For others, the combination may be far less advantageous. In addition to considering the "receivers' view" of the discussion, research also needs to address which approach works best for which supervisors and organizations.

H. Ideas for Improving the Process

We turn now from the effects of the appraisal process to ideas advanced for improving it. The literature on employee reactions to performance appraisal suggests several ways to improve receptivity to the process. Steers and Lee (1983) summarize six conditions from the literature on organizational psychology and management that facilitate performance appraisal systems in business and industry:

1. Extensive and frequent communication.
2. Widespread belief in the integrity and legitimacy of the particular appraisal system.
3. The extent to which a system is considered by employees to be instrumental for the receipt of desired rewards.
4. Highly specified and clear job standards and expectations.
5. Employee participation in the design and implementation of the appraisal system.
6. Clear support from the top management, who abide by the results.

Participatory performance appraisal is broadly supported in the literature (e.g., Steel, 1985; Wexley, Singh, & Yukl, 1973), with stronger worker motivation, high rates of satisfaction, and high productivity associated with participatory formats.

The manner in which supervisors handle the interview in which feedback on performance is provided influences subordinates' reactions to the interview and their motivation to do something about their subsequent job performance (Burke, Weitzel, & Weir, 1980). Taking enough time to talk about ideas and feelings and looking forward rather than backward have been shown to increase the effectiveness of performance reviews.

There is also an extensive literature on the effectiveness of rater-training in improving the outcomes of performance appraisal. Training efforts address individuals' limited information processing capabilities as well as their susceptibility to situational distractions and personal biases. For example, in an effort to reduce halo bias, various rater-training techniques have been developed. A key assumption is that increased observation of performance-relevant behavior will reduce halo error. Feldman (1983) has pointed out that training raters to recognize relevant behaviors should improve accuracy. Williams et al. (1985) suggest that such training should be extended one step further—raters should receive instructions on how to search for relevant behavior when it is not present. Obviously, training in all instances must attend to the search biases inherent in person perception processes.

Rather than focus on rating errors, Bernardin and Pence (1980) advocate that "rater training should concentrate on enhancing the accuracy of ratings through discussion of the multidimensionality of work performance, the importance of fair and critical evaluation, and the development of stereotypes of effective and ineffective workers" (p. 61).

Numerous researchers (e.g. Landy & Farr, 1980) have examined the effects of various rating instruments and formats on accuracy in performance appraisal and have found that there seems to be no sure way to solve the problem. One format that promised resistance to rating errors was the behaviorally anchored rating scale (BARS). But numerous studies evaluating the BARS format have not borne this out (cf. review by Jacobs, Kafry, & Zedeck, 1980). Comparisons across BARS and behavioral observation scales (BOS), in conjunction with rater training relevant to the scale, indicate that no single training procedure or combination of training procedures is consistently best (Pulakos, 1984). Banks and Roberson (1985) argue that "without built-in controls for rater biases, idiosyncrasies, and inconsistencies, contaminants are likely to affect appraisal judgments and, hence outcomes *no matter how many elegant psychometric properties a format possesses*" (p. 139).

Based on their research on the performance appraisal of nurses in four different hospitals, Zammuto, London, & Rowland (1982) argue against using a standardized appraisal instrument to compare performance across institutions. Such an approach does not recognize the environmental context in which ratings are applied. Zammuto et al. also argue for the use of multiple rater groups in the appraisal process.

Recognizing the problems with rater-bias and the various appraisal formats, many organizations use multiple measures of performance, adding systematic colleague and self-evaluation to the review process. But caution is in order here, too, as each of these sources possesses common and unique problems as well as certain advantages.

The basic problem with self-appraisals is that a number of studies have shown that employees consistently rate themselves higher than they are rated by their supervisors or peers (Holzbach, 1978; Thornton, 1980). Meyer (1980) reports that when asked to rate their own job performances, 40 percent of employees in jobs of all types placed themselves in the top 10 percent, while virtually all remaining employees rated themselves either in the top 25 percent or at least the top 50 percent, that is, "above average" (p. 293). Self-evaluations have been found to be particularly lenient among professionals. In one group of higher-level managerial and professional employees, over 80 percent placed themselves in the "top 10 percent" category (Meyer, 1980). Shapiro and Dessler (1985), in a study of 240 supervisors in a Southeastern hospital, had findings consistent with Meyer's.

Faculty have also been found to overrate themselves, especially on their teaching performance. In a study of faculty in 24 institutions—eight community colleges, eight liberal arts colleges (four of which were selective), and eight universities (four of which were research focused) (Blackburn, Boberg, O'Connell, & Pellino, 1980), 90 percent of the faculty judged themselves above average or superior teachers. Gaff (n.d.) had 99 percent in the same two categories. Furthermore, these faculty believe that their students will rate them accordingly, but not their peers.

Unlike self-appraisals, peer appraisals have been shown to meet acceptable standards of reliability. It is suggested that this high reliability is probably a function of the daily interaction among peers and the use of multiple raters, which provides a more reliable measure than a single rating (e.g., Love, 1981). As for problems with the use of peer ratings, De Nisi, Randolph, and Blencoe (1983) report that negative peer rating feedback lowers group cohesiveness and group interaction and performance on subsequent tasks.

Mount (1984) looked at the interaction of multiple ratings in a high technology corporation in the Midwest. He found a great deal of similarity among the ratings obtained from three sources—supervisor, self, and peers and subordinates—and advocates a performance appraisal system that draws information from multiple sources.

Zey-Ferrell and Ervin (1985) document the need for peer involvement because of the high regard faculty have for that source of evaluation. The use of peers could temper the faculty complaint that students are the exclusive judges of their teaching. At the same time, studies show that faculty participation in the process is rare, despite the theoretical reasons for their doing so. In practice the costs in both time and in psychological discomfort may exceed the benefits (Centra, 1975; Doyle, 1983). When correlations between student and peer ratings of faculty teaching are conducted, they run reasonably high (around 0.65) (Blackburn & Clark, 1975).

Another problem of peer rating, however, enters when the findings on administrator ratings of faculty and faculty self-ratings are examined. The correlations between division heads' and deans' ratings with those of students are about .4, but between the individual professor and the other three constituencies they fluctuate around .0. When faculty judge their own teaching—and they do constantly—how they rate themselves has little relationship to how students, their peers, or administrators rate them (Blackburn & Clark, 1975; Centra, 1972). Doyle and Webber (1978) and Marsh, Overall, and Kessler (1979) found closer agreement between students' ratings and self-ratings.

I. Conclusions of the Literature Search

Table 2, which follows, provides a summary of the major research findings on the important elements in performance appraisal. Drawing on these findings, and taking

TABLE 2

Research Findings on Important Elements in Performance Appraisal

ELEMENT	FINDINGS/ISSUES	STUDIES
Kinds of rater errors	Stereotyping of individuals can produce false performance evaluations.	Feldman (1981) Dipboye (1985)
	Appraisal rating errors due to halo or systematic distortion distort the rating process.	Borman (1979, 1983) Nisbett & Wilson (1977) Bernardin & Pence (1980)
Influences on rater's judgments	The purpose of the performance appraisal affects the way raters collect, combine, and recall information.	Hamilton, Katz, & Leirer (1980) Zedeck & Cascio (1982) Williams, DeNisi, Blencoe, & Cafferty (1985)
	Initial impressions lead to observation bias.	Balzer (1986) Hogan (1987)
	Prior commitments influence subsequent judgments.	Staw (1976, 1981) Bazerman, Beekun, & Schoorman (1982)
	The race and sex of both the rater and ratee can influence the outcome of the rating process.	Taylor & Fiske (1978) Schmitt & Lappin (1980) Hamner, Kim, Baird, & Bigoness (1974) Maruyama & Miller (1980) Deaux & Emswiller (1974) Rose (1978) Wexley & Pulakos (1982) Abramson, Goldberg, Greenbaum, & Abramson (1978)
	Physical attractiveness influences performance ratings.	Heilman & Stopeck (1985)
	The age of the ratee affects appraisal outcomes.	Cleveland & Landy (1981) Rosen, Jerdee, & Lunn (1981) Schwab & Heneman (1978)
Effect of the performance appraisal process on the ratee	Low ratings can cause a drop in employee morale and commitment.	Pearce & Porter (1986)
	The performance appraisal process does not have a positive impact on subsequent job performance.	Pearce & Porter (1986) Dorfman, Stephan, & Loveland (1986) Bernardin & Beatty (1984)
	Perceptions of fairness and accuracy in performance appraisal depend heavily on trust in the supervisor.	Fulk, Brief, & Barr (1985)
	Employee response to performance appraisal depends on the characteristics of the system itself.	Landy, Barnes, & Murphy (1978) Landy, Barnes-Farrell, & Cleveland (1980)
	Favorability of feedback affects its perceived accuracy.	Stone & Stone (1984) Shrauger & Schoeneman (1979) Dipboye & de Pontbriand (1981) Shields (1984)
	Ratee participation in the performance appraisal process affects satisfaction with it.	Mount (1983) Dipboye & de Pontbriand (1981) Steel (1985) Wexley, Singh, & Yukl (1973)
	The ambience of the feedback setting affects the ratees' response to the feedback.	Ilgen, Peterson, Martin, & Boeschen (1981) Burke, Weitzel, & Weir (1980)

TABLE 2-*Continued*

Research Findings on Important Elements in Performance Appraisal

ELEMENT	FINDINGS/ISSUES	STUDIES
Effect of the performance appraisal process on the ratee *(continued)*	Performance-contingent (merit) reward systems do not always improve individual or organizational performance.	Pearce, Stevenson, & Perry (1985) Watts (1985)
Nature of the performance appraisal process	Salary decision making and individual development components of performance appraisal need not be separated.	Ilgen & Feldman (1983) Prince & Lawler (1986) Dorfman, Stephan, & Loveland (1986)
	Goal setting and performance counseling should be separate from performance appraisal.	Finn & Fontaine (1984)
Rating instruments and formats	Despite the numerous attempts to develop error-free approaches to performance appraisal, none currently exist.	Landy & Farr (1980) Jacobs, Kafry, & Zedeck (1980)
	Self-appraisals usually exaggerate toward the positive end of the rating scale.	Holzbach (1978) Thornton (1980) Meyer (1980) Shapiro & Dessler (1985) Blackburn, Boberg, O'Connell, & Pellino (1980)
	Peer appraisals meet acceptable standards of reliability.	Love (1981) DeNisi, Randolph, & Blencoe (1983)

into consideration the theoretical propositions on the nature and effects of performance appraisal, 15 summary statements emerge:

1. The literature shows that most current systems of performance appraisal or evaluation do not lead to improved performance. It does show, however, that performance appraisals *can be dysfunctional*, lead to reduced productivity, and create morale problems. The outcomes of performance appraisal have a significant, often negative impact on the climate of the organization and the commitment of its employees.

2. Colleges and universities have not paid enough attention to the scientific theory behind performance evaluation systems, to the research on their effects, or to unique characteristics of the college or university environment.

3. Theory and research on motivation, the way in which motivation contributes to performance, and how it is appraised show that the sequence is *multifaceted* and *individualized*. Therefore, to be effective and fair, performance appraisal has to be individualized. A standardized approach and standardized appraisal forms are bound to discriminate against some workers.

4. Rater and ratee characteristics, the measurement vehicle, and the context in which the rating is done are all important factors related to the validity of performance evaluation.

5. Of the several conditions shown to facilitate performance evaluation systems, employee participation in the design and implementation of the appraisal system is one of the most important.

6. Performance evaluation is intended to serve both developmental and administrative purposes. The issue of whether performance feedback and counseling (the develop-mental function) is most effective when it is separated *in time* from

the merit increase discussion (the administrative function) is unresolved. It appears that the functions need to be separated for some individuals, but not for others.

7. While we did not find a significant body of research addressing the interrelationship, we conclude that personality attributes are related to how and how often appraisal is desired and received.

8. Research has shown that feedback indicating that one is "satisfactory" is disconfirming for many individuals.

9. Untrained appraisers can transmit messages that are received other than as they were intended. Research has shown that rater training can improve the process of performance evaluation, both in terms of reduced errors and rater and ratee satisfaction with the process.

10. Performance-based reward systems usually produce the behavior that is desired and rewarded, but they often produce other unwanted outcomes, such as an excessive focus on individual rather than collective performance, emphasis on short-term versus long-term results, and greater concern for extrinsic rather than intrinsic rewards.

11. Unless performance-based reward systems are carefully structured, the quantity of activity rather than the quality of the outcome will take precedence.

12. The nature and conditioning of professionals—especially faculty—build in a predictable conflict between their autonomy and the organization's goals. Performance evaluation is perhaps the best example of this conflict.

13. Not all workers want the same kind of reward. Whereas some respond to monetary incentives, others prefer better working conditions, more variety or responsibility in their assignments, public recognition of their accomplishments, etc. Individual preferences for intrinsic or extrinsic rewards are not recognized by most performance-based reward systems.

14. Research on job enrichment and on the social context in which behavior occurs offers some useful insights into the motivation-behavior sequence.

15. More research is needed on how evaluation apprehension, the task itself, and social cues affect motivation and performance.

III. How the Findings Fit with College and University Structures, Norms, Values, and Practices

While learning what theory predicts and what propositions about performance appraisal of professionals are supported by research, we noted that most of the evidence comes from workers in organizations other than colleges and universities. The next question, then, is: What should one expect if these principles are applied to faculty in their work setting? Another way of putting the question is to ask: In what ways and to what extent are typical higher education structures, norms, values, and practices in accord with the evidence? This section, which departs from the form of a literature review, offers an overview of the key considerations that influence the conduct of performance appraisal in the college or university setting. For those in postsecondary education this section is merely a reminder of well-known conditions and not a demonstration of their existence. Consequently there are relatively few citations to document these statements.

A. Structures

While the typical college or university organization chart shows a hierarchical, pyramid structure with *a* president, *x* deans, *y* (> *x*) department chairs, and still more faculty; in most instances chairs have not only been selected by the faculty but are in their administrative role for a limited time period.[1] The majority of chairs expect, and are expected by the faculty, to return to faculty status. (This expectation also applies to deans, vice-presidents, and even the president, but probably occurs less frequently at each higher level.) The structure, consequently, does not place the immediate supervisor who will be responsible for performance appraisal in the same relative position of authority that exists in a business organization. The chair is first a representative of the faculty to those higher up who control the institution's resources, not a boss. In a study of faculty and (non-academic) administrator stress in a research university, Blackburn, Horowitz, Edington, & Klos (1985) found that the principal source of stress for the administrators was their immediate supervisor but for faculty it was the lack of time to accomplish tasks they had set for themselves. Faculty have little trouble with their boss; many do not even think they have one.

A second structural difference between colleges and universities on the one hand and business and industry on the other is the successful career path. In the latter, the pyramid clearly is in place and, for all but a handful, worker advancement includes many promotions along the way but comes to a halt before the top of the ladder, salary may continue to increase for those professionals until retirement, but position and title (and the salary that accompanies the higher office) will not. For faculty, however, it is expected that every successful assistant professor will not only be promoted to associate professor but also will become a full professor. That is, in academe everyone can make it to the top; even institutions where promotions are frozen from fear of over-tenure still subscribe to the principle of awarding all good work with the highest rank. In some ways, however, and for some individuals, colleges and universities have the pyramid structure. For many outsiders, success in academe is viewed as climbing up the administrative ladder, that is, becoming president (and of a distinguished institution). Indeed, some faculty aspire to administrative positions from a desire to move the organization in certain new directions, for power, for money—for any number of reasons. But they usually did not become a faculty member to become a dean; something happened along the way. Furthermore, for those on the inside, that is, for faculty, success is *not* becoming an administrator. Faculty who do are regarded poorly by their former colleagues, just as doctors judge each others' medical reputations rather than their abilities as administrators.

---

[1] In some colleges and universities the chair is a "head," appointed by central administration for an indefinite (but implied continuous) period. These individuals are clearly in the administration and not the faculty. We discuss the case of unionized faculty later.

A virtue of the faculty career path is that making long-range plans and having long-range goals is realistic. A limitation to the academic career path is the few recognitions of advancement—just two, in fact, and for some faculty (e.g., in community colleges with no rank other than instructor), none at all.

These structural differences obviously affect both the consequences of a performance appraisal and the way in which it is likely to be carried out. When a chair and a faculty member are likely to reverse roles the following year, the content and process of evaluation can be expected to be different from situations where the rater-ratee relationship is a long-standing one.

B. Norms and Customs

"Assigned" duties occupy but a small fraction of the typical faculty member's work week. Faculty are assigned *x* number of hours in the classroom (typically 15 in a community college to six or fewer in a research university) (they may be assigned what to teach, but not how to teach it), two to five office hours for advising and counseling, a committee assignment or two taking a couple of hours a week, and an occasional ceremonial event to attend. After that, faculty have great freedom to pursue whatever academic activities they wish. In unionized colleges and universities the conditions of work are likely to be highly specified, but even then there are more unspecified hours than assigned ones.

Faculty are not 9-to-5 employees, working on the same schedule and in the same general work space all week long. To appraise their productivity and their effectiveness for the organization and to develop a system that would fairly compare Professor A perfectly to Assistant Professor B is no small task.

Faculty are, of course, employed by an organization but they view themselves as independent professionals. Real professionals, faculty agree, are *not* supervised. They know what is expected of a person in their role and see no need for organizational functionaries to look over their shoulder. Professionals—be they faculty, doctors, or lawyers—earn a reputation for the quality of their work, not from a number on a rating form. Quantitative rating forms are anathema to most professionals.

A third norm is long work hours—as many as 55 hours a week for university faculty, as confirmed by several studies (Ladd, 1979). A good faculty is a collection of high achievers who are self-motivated and responsive to intrinsic rewards. These norms conflict with a merit pay system (Staw, 1984), as Deci (1975) and McClelland (1961) found when incentive plans were put before high achievers. Merit pay, or merit increase in salary, is defined in several ways. We define it as that portion of a salary increase an individual receives above the percentage (or amount) all receive as "cost of living" or adjustments for past short-falls. The range of salary increases based on merit is from none (a zero percent raise is a reality even when the average is six, seven, or any percent) to complete. Finkelstein (1984), citing Ladd (1979), argues that faculty subscribe to differential pay as determined by varying levels of performance. His finding may hold for faculty in research universities but not for the vast majority of faculty elsewhere.

Another norm affecting a performance appraisal is that of peer review. Faculty contend that the best—more likely, the only—qualified judges of their performance are those colleagues who have the same expertise as they do. Peer judgment takes precedence over both supervisor (head or chair) and client (student). Self-appraisal is also important and is discussed later. A performance appraisal from a chair or head may carry authority and affect salary. It will probably not, however, be acted upon unless it is based on peer judgment. In addition, as we learned in our case studies, a rating form will be discredited and its inadequacies endorsed by peers. Therefore, an effective performance appraisal plan will have peers involved, as they are in promotion and tenure decisions.

A fifth norm is an unfortunate one, if performance appraisal is going to be linked with salary. The institution may announce that salary is to be determined solely by performance (say publications in a research university), but the truth of the matter is that salaries in some institutions are related to having been a chair or a dean (Kasten, 1984). When a faculty member takes on an administrative role, salary escalates. In those institutions where it does not drop down to a corresponding level when the person returns to faculty status (as it does in many institutions), performance appraisal is not likely to resolve the salary inequities. This fact needs to be realized before a performance appraisal plan is instituted.

While faculty life is typically described in terms of the three roles of teaching, research, and service, seldom are the three accorded equal value on any one campus. Community colleges and liberal arts colleges are more likely to give the greatest weight to teaching. At these institutions research may be seen as an added bonus (but not a necessity); in some, research may even be seen as a detraction from attention that should be given to students; and in still other "teaching" institutions, research may be viewed as a requisite for advancement. In every instance, institutional norms will affect a performance appraisal plan.

A seventh norm is the influence of the marketplace that causes salaries to differ across academic disciplines within the same institution. Faculty in scarce fields (today, for example, Ph.D.s in accounting or computer science) are going to be paid more than their colleagues in the liberal arts even when the latter may be doing more work. Equally meritorious work will not necessarily receive equal merit pay. The salary differences across disciplines will be least in principally undergraduate institutions and most unionized faculties, and greatest in research universities where the external marketplace, including peer institutions, sets the salary structure.

C. Values and Beliefs

Just as special norms govern the life of faculty and other professionals, so do different kinds of values permeate their cultures. In the social/psychological category, faculty value fairness and equity although they sometimes confuse "fair" with "equal." There are to be no favorites of the administration; all are to have equal access to rewards. Intrinsic rewards, challenge, and job enrichment (learning new skills or knowledge) are also high on faculty members' value lists. These priorities can be expected to change over the course of a career as faculty respond to a changing environment and their own growth and development.

Academics also believe in "openness," their "right" to know about decisions that have been made. They believe each faculty member should carry a full load of responsibilities. They seem to be opposed to a division of labor where, say, some faculty would only teach and some would only conduct research, even though that would be the preference of both sets of individuals. In addition, faculty value congeniality, or, at least, minimal inter-personal conflicts.

Faculty also subscribe to other values held by professionals—no sacrifice of standards, the responsibility to police themselves, the primacy of the client, and setting the standards for admission to the profession (which typically include expertise in an esoteric body of knowledge). In addition, professors like successful students, ideas, and learning simply for its own sake. They frequently characterize themselves as perpetual students.

Professors also believe they are very good at what they do. Almost all rate themselves above average and the majority say their teaching is in the top ten percent, a "fact" that they are certain students will confirm (Blackburn, Boberg, O'Connell, & Pellino, 1980).

They expect these same basic values to be held by academic administrators. That most administrators also value productivity, efficiency, order, and organization puts them at odds with faculty, for these values come in conflict with freedom, self-policing, and the

generation of knowledge. Having profound insights and discovering new truths are not events that can be put on a timeline, so faculty assert.

As is apparent, these values in part contradict what would be best for faculty growth and development—their self-efficacy. Fairness would have everyone evaluated against the same set of standards and the results executed in the same manner. Openness means learning that others are judged to be better than they, not something to which faculty will take kindly or react positively. Fairness could mean sacrificing the freedom to allocate one's efforts as he or she desires since "good scores" on a single standard determine "merit." Subjecting all faculty to the same performance appraisal may disadvantage some simply because of personality characteristics, which are not to be judged if universal standards of excellence are to prevail. (We return to the dilemmas created here after we look at practices.) Faculty value-systems often conflict with administrative processes such as performance appraisal.

D. Practices

While many of the structures, norms, and values are similar across postsecondary institutions, how performance appraisal is practiced certainly is not. Although the frequency distribution with which different procedures are used is not known, we know there is great variation.

First of all, practices are in part determined by conditions in the external world that must be taken into consideration. Legislators or boards of trustees can call for accountability and the institution has to respond. Hard times mean economic constraints, freezes, and cutbacks. Personnel appraisal practices can change. Such practices are related directly to the salary policy in effect in a given institution. In those colleges and universities where annual salary increases are allocated evenly "across the board," performance appraisal is likely to be less rigorously pursued. But in those colleges and universities where merit is the primary determinant of annual salary adjustments, performance appraisal takes on much greater importance. Keeping these unpredictable forces in mind, we turn to the kinds of practices found throughout higher education in this country.

The diversity of salary-increase practices extends across the entire spectrum—from no merit (i.e., identical across-the-board raises for all) to almost every combination (including tenths of a percentage point) of across-the-board (sometimes called cost-of-living), with some additional percentage set aside for merit, to complete merit systems, which mean that a zero percent raise is both an option and a fact. In addition, there may be bonus dollars (a one-time recognition that does not get added to the base salary) or a special pool the dean or vice-president can use outside of the general agreement for the following year.

As for performance appraisal, it too spans the full spectrum, from absolutely none to an annual complete portfolio. (Both of these exist in community colleges. See Poole and Dellow, 1983.) In the cases of fairly regular appraisal, that is, in those instances beyond tenure and promotion decisions, the level at which it takes place also extends from the chair to the presidential level.

These assertions are just as true for faculty who are unionized as they are for those who are not. They are also independent of the bargaining agent—independent, AAUP, national, or a statewide system. The chairs may or may not be in the bargaining unit, their terms may be fixed or indefinite, and their role may be highly specified or not. There are instances of unionized faculty having a greater role in performance appraisal and voice in selecting supervisors than there are in non-unionized faculty, and vice versa. In short, collective bargaining per se is not necessarily related to performance appraisal practice in any institutional type.

At one end of an informal/formal continuum is what might be labeled the laissez-faire performance appraisal. When salary decisions must be made, the personnel

committee holds a private, confidential session. The president or dean attends, but may only listen, not speak. The roster of faculty is then read, one by one and each committee member comments on the merits of each individual.[2] There are no folders or files available on the staff. Some discussion takes place if there are significant discrepancies between ac-counts of how well Jones did this year. A couple of weeks later each faculty member will receive a letter expressing appreciation for the contributions he or she made during the year and lamenting the institution's financial situation. The letter then states that the individual's salary for next year is *z* thousand dollars.

Professors can calculate their salary increase, but they will not know if it is above or below the average (nor will the executive committee that discussed their strengths and weaknesses). The performance appraisal is simply an uninformative number, a salary figure.

Next on the continuum are practices that provide a performance appraisal. For example, student evaluations of teaching are now quite common. In some unionized institutions, student evaluation of faculty teaching is mandated by the collective bargaining agreement. These evaluations may involve scale scores on performance items and often norms for an administratively established peer group. Sometimes only the faculty member receives student evaluations and chooses whether or not to submit them to the dean for merit-raise consideration. Sometimes the scored forms are also sent to the supervising administrator. (In one liberal arts college we worked with, student course evaluation is the sole determiner of merit increase.)

Other practices for evaluating teaching include visits by senior professors, visits by the chair (who may also use a rating form), conferences with the chair after a visitation, and the sharing of the chair's evaluation, including the signing by the faculty member that she or he has seen the appraisal. In some colleges and universities, teaching is appraised by additional means. Syllabi are examined, student papers are read, examinations are looked at. We found few instances of the more extensive practice.

In our case studies we learned that student evaluations had inconsequential effects on faculty teaching behavior. The few who did report making some change in their teaching as a result of student evaluations made principally cosmetic alterations. One became more stringent with late arrivers; another substituted one reading for an existing one.

For a number of reasons, faculty make few changes as a consequence of student appraisal. First of all, most receive high scores. "Why change?," they ask. Second, and more important, almost without exception faculty give less than full credence to student evaluations, and some consider them worthless. They all know a faculty member who is not as good as they are (is principally an entertainer and/or is not up-to-date) and who gets higher student ratings than they do. In all, student evaluations as a meaningful form of performance appraisal receive a very low grade from faculty in the majority of institutions.

Turning to scholarship, for those who are engaged in traditional research (publishing articles in the prestigious journals, having one's sculpture purchased by a distinguished museum, having a history accepted by a university press, obtaining a competitive grant), performance appraisal seems to be working well. The peer review process—the critique and revisions made before publication—are appraisals from which one can learn and grow and are accepted by both faculty and administrators as right and proper. Furthermore, although not with universal unanimity, relative worth can be assessed so that merit can be more fairly awarded on the basis of these criteria.

The problem with the accepted practice, however, is that it affects only a small share of the professoriate. Scholarship, or creative work, for the vast majority of the faculty is

[2]This discussion is for merit pay for the following year, not for promotion or tenure decisions, which are likely to have a more formal process, even for institutions at this end of the continuum.

frequently not subject to peer review. Presenting a paper at a regional conference, creating a new course, converting an old course to a self-paced machine program available in the library, developing educational software, preparing a special lecture for a colleague's class—there is a long list of creative activities that most colleges value and need but seldom appraise.

As for service, the most common assessment of it is "no assessment." Faculty report at the end of the year what committees they were on, science fairs they refereed, high schools they visited to recruit students, programs they staged on campus to attract students, supervision of student organizations—again the list is long. But not a single faculty member we interviewed said anyone assessed how well they did in any of these activities. They agreed that doing them might affect their next year's salary, but most put it in negative terms. If you did not do any service, you could get penalized; but doing some pays little, if anything.

At the same time, both faculty and administrators depend on the internal service activities of faculty. Faculty strongly want to have a voice in institutional matters and realize they will not if they do not participate in governance-related activities.

For many institutions, faculty performance appraisal is carried out at the departmental level where a salary review committee develops merit increase recommendations based on the annual activity reports of all departmental members. While this approach looks at the full range of an individual's efforts and involves peer review, it is very often performance appraisal with no feedback.

Last, and at the other end of the scale from no systematic appraisal through partial appraisal, we come to the highly structured procedure we found in two institutions and which we know to be used at several others. In this quantified approach to performance appraisal, an intricate set of activities was detailed into almost every conceivable kind of performance, with weights assigned to each. A computer spreadsheet was necessary to lay out its details. An elected personnel committee spends hours each year scoring each person on each scale, weighting by effort expended (self-reported), and coming up with *a* number for each member of the department (see Tucker, 1984, pp. 153-162). (Miller, 1974, 1987, and Seldin, 1984, also have sample forms, and they call attention to individualization as well.) An individual's number then is compared with everyone else's, and where that number is on the score sheet determines what fraction of the amount or percentage set aside for merit each would receive. While such a system is technically "fair," it does not constitute a developmental form of performance appraisal. Individuals learn their number, but little more. As a performance appraisal, it probably is no more effective for changing individual behavior (except in undesired ways) than the laissez-faire technique we described at the other end of the scale.

What can one infer and conclude from these practices? Performance appraisal of faculty does not seem to be done well even in those places that do collect some evidence of performance. In every instance, the procedure violates the principle of separating appraisal for the sake of growth and development from using it to award merit pay for good work. These performance appraisal systems assume dollars to be the only incentive to which faculty will respond.[3] Yet we know from many studies (Blackburn, 1974, 1980) that faculty are not primarily the economist's rational person. Most faculty care equally or more about work conditions—the quality of their colleagues, the kinds of students they have, the holdings of the library, the courses they will teach, the reputation of the department—than they do about dollars, as important as a decent income is to them.

[3] Lawrence (1988) analyzes the indefensibility of the proposition that if you pay for good teaching, you will get better teaching. She also dissects the companion proposition so often advanced by those who know how to solve the "teaching problem," namely, have the graduate schools ensure that all Ph.D.s are competent teachers.

Second, many of the appraisal procedures used by colleges and universities judge everyone on the same criteria regardless of the organization's different requirements and its need for different kinds of individuals. Often the quantity of the work that is done, not the kind or quality of work, is the overwhelming factor.

Third, colleges and universities are pitting colleagues against one another in a zero sum game. For every professor who gets more, another gets less. The student evaluations play colleague against colleague. Half are judged to be performing below the median, even when the lower scores are still more than acceptable.

Fourth, the process is probably dysfunctional. The haves and the have-nots grow farther apart each year, and few performance improvements are realized. Social support goes down, along with morale. Alienation exacerbates the problem. When this happens, the institution and the individual faculty are less likely to accomplish their goals.

IV. A Recommended Performance Appraisal Plan

We begin with some simplifying assumptions about performance appraisal then add the factors that complicate matters. One assumption we make is that what is good for the faculty member is good for the institution. That is, we want to develop an appraisal plan that faculty will endorse and support because it will help them better accomplish their professional goals. At this stage, then, the plan will have no relationship to external rewards, such as merit salary increases or the achievement of institutional objectives. In Section V we address feasibility questions—the conditions that would have to prevail for our proposal to be incorporated in a college or university and what alternatives to our proposal would still constitute positive steps forward.

A second assumption comes from the evidence we reviewed in Section II. One conclusion from our studies is that to be effective, performance appraisal must be individualized. A uniform system discriminates against some individuals. We want a plan that will work for everyone. Assuming the conclusions found to be true for workers in other kinds of organizations are valid for the performance appraisal of faculty, then a plan sensitive to individual variation needs to be developed.

With these assumptions we can begin to analyze what individual factors need to be taken into consideration when appraising faculty in their various roles. We will see immediately that discipline (including sub-specialty) and place of work will affect every appraisal plan. We start with these factors as they relate to the scholarly/creative faculty role.

The generally accepted meaning of scholarship today, namely, publications, differs dramatically across disciplines in the kinds of products and their rate of output. A biologist with an ongoing research program will have a laboratory producing brief, multi-authored abstracts and articles at a high annual frequency. On the other hand, an historian writes books or monographs, most often alone and certainly not at a rate of more than one per year.

Even within a discipline great variation is the rule, not the exception. An experimental psychologist can be expected to be a frequent article producer, whereas a developmental psychologist depending on longitudinal data more likely will publish monographs at greater intervals of time. Quite clearly, then, a uniform rating scale, even when weighted for type and "quality" (the status of the journal or the reputation of the press publishing the book) will discriminate against some faculty and will not accomplish the aim of enhancing distinguished scholarship.

The problem of measuring scholarly products compounds when the place of work is introduced. The publish-or-perish syndrome is typical in research university settings. Research universities, however, constitute but a small fraction of the country's higher education institutions. The scholarly/creative component of the majority of professors in this country is diffuse. It is not that most are unscholarly; rather, they present their creative efforts in a wide variety of ways, almost none of which possess a standard with which its quality can be judged or at a rate that can be advanced as "normal."

Faculty in community colleges, liberal arts colleges, and regional universities also are creative individuals who take pleasure in solving problems. Most frequently the problems they work on are teaching related—the design of a new lab experiment, converting a course to individualized instruction, creating an interdisciplinary course, establishing an alternative curriculum, finding a way to get an underprepared student able to understand and do mathematics.[4]

Institutions recognize these creative contributions, even though most are invisible to colleagues elsewhere. Consequently, activities of this kind do not receive the

[4]This is not to say that these faculty members' creative work cannot be of the traditional kind as well, just as research university professors also engage in creative teaching-related problems.

professional review that an article submitted to a scholarly journal receives. Clearly appraisal of such creative work calls for an individualized approach. The problem is how to assess the quality of the creation. (See Section V for possible ways of assessing the quality of creative work.) We see that place of work is going to be related to the kind of creative work in which faculty typically engage.

Turning to the teaching role, disciplines sometimes have distinctive instructional styles (case method in law and business schools, the laboratories in some of the sciences, etc.) but are essentially independent of setting. U.S. history is taught in much the same lecture manner in all institutions, although the aims of the course, the amount and nature of the assignments, and the kinds of examinations can differ appreciably from place to place. The individualization of teaching can vary by the type of course—whether it is a "regular" class, a remedial one, team taught, a laboratory, or a massive lecture—and this needs to be taken into consideration.

Individualization in teaching is threatened by the typical way it is assessed, namely, by student evaluations. Even when student rating forms are designed to be appropriate for a variety of conditions by allowing faculty to select from a pool of statements those that will be used in assessing them, the forms still discriminate because they establish a uniform set of standards and assume that certain behaviors are good, and that the absence of these behaviors constitutes proof of poor teaching. For example, almost all forms ask if the instructor is well organized. Teachers who believe that learning is more effective when students have to create order than when the instructor provides it for them may be penalized by most student rating forms.

In addition, we know that student rating forms as performance appraisal instruments do not accomplish what is desired, namely, the improvement of instruction. As noted in Section III, faculty have good reasons for not changing their teaching behavior on the basis of information received on student forms. Something more is needed if performance appraisal of teaching is to accomplish its goals.[5] Among other things, it needs to be individualized.

The service role is seldom appraised. As we have noted, faculty tend to report annually the service they have engaged in, but how effective they were in any of these activities is not often judged. Some institutions place a high value on certain kinds of service (student advising in community colleges, for example) and others give the same service essentially no weight at all (research universities). One is expected to do a full share of this kind of activity, but few see professional growth associated with it. In fact, more often than not such assignments are judged as dysfunctional because they take time away from teaching and creative work. Performance appraisal of this role is meaningful for those work places that genuinely value service. In the former settings, evidence will be needed and the performance standards individualized.

Two questions remain: What form might meaningful performance appraisal for faculty take? How should its outcome be related to differential rewards?

We believe a portfolio system would accomplish the goal of continuous faculty growth and development, the realization of the individual's full potential. We believe that the portfolio process fits the conclusions from the research literature and college and university norms, values, structures, and practices. We know the idea is not novel. In higher education it comes from growth contracts such as those developed at Gordon

[5] When promotion and tenure decisions are being made, that is, when external rewards are at stake, the organization introduces other kinds of evidence of teaching quality—visitations (which occur in some institutions on a regular basis), examination of syllabi, examinations, student success, etc. But typically an in-depth review happens only twice in the person's career—at promotion to associate professor (which includes tenure) and to full professor. Performance appraisal that will help faculty develop professionally needs to take place more frequently than twice in a lifetime.

College.[6] In business and industry, performance contracts are often a component of management-by-objective systems.

By definition, the portfolio process is individualized, and thus in accord with our foremost research finding. The process can also reduce gender and racial bias. As Rossi (1980) has learned—with respect to the differing, changing needs of men and women at different ages (women becoming less affiliative and more competitive at successive ages, and men just the opposite)—goals may change and in different ways. There is no equivalent evidence on racial bias, but certainly the interactions of faculty members with peers and chairs are going to be more influential in eradicating these biases.

The portfolio process can directly address the discipline and specialty differences on kinds of products and rates of production. Different expectations can be held that eliminate the inequities of rating scales and the production of trivial items just to meet a uniform quota. In this way, the portfolio process is well suited for setting long-range as well as short-term individual goals.

In addition, the process obviously involves the individual being evaluated and certainly could involve peers. (Our specific plan does.) Multiple information sources, including both self and colleagues, have been demonstrated to be related to successful performance appraisal systems. With most current higher education practices, the only appraisal is from clients (students on teaching forms), a practice engaged in with no other professional group. (Not even relatively low-status professionals, such as floor nurses, are graded by their patients.) A sound portfolio process involves the affected person and has the peer judgment faculty respect, both of which are in accord with the research findings and with faculty values.

The portfolio process also allows for the separation of attention to growth and development from merit and salary increases, or can for those faculty who would prefer the two aspects to be considered separately. Certainly the evidence collected for the portfolio is going to be used to determine rewards, but the two acts do not have to be simultaneous for those who believe talk of dollars will contaminate discussion of career plans and trajectories. Furthermore, in a portfolio setting, rewards other than dollars are a natural topic of discussion. The process makes it possible for faculty to make known what else they desire—space, a piece of equipment, a block of time, an altered schedule, the need for professional contact—a list is easy to generate.

In addition, we believe the portfolio process will be good for the organization as well. In a national survey of chief academic officers in 210 community colleges, performance appraisal was highly endorsed (although considered expensive). It was used principally for faculty development and consequently well received by the faculty (Renz, 1984). When the chair knows with some regularity the wishes and accomplishments of all staff, departmental planning can take place. Institutional goals can be set with the knowledge that the staff can accomplish them. As departments move into the process, goals can be accumulated for the institution. Portfolios ought to improve the organization as well as benefit the individuals.

What might a portfolio plan look like and how might one work? By way of illustration, faculty members would assemble those artifacts of their lives that demonstrate where they have been and what has been accomplished. The faculty member would then draw up a plan of where he or she wants to go in both the short run and over the long haul, both being subject to alteration, especially the latter, as time passes. Next, the faculty member would delineate what is needed to get from her or his current stage to the next one and what will constitute evidence that the new stage has been reached.

[6] The idea has not spread very far or very fast. The reasons for its slow adoption are suggested in Section V. Bare (1979, p. 15) also mentions New College of the University of Alabama, Ottawa University, William Jewell College, and Mercer County College as having experience with growth contracting.

The portfolio would then be shared with some colleagues and the institution (chair, dean, whoever is appropriate). Questions would be asked to clarify means and ends, opinions would be offered as to the feasibility of the plan and whether or not it is realistic. The institution would make a commitment to what it can provide. It would also tell what it believes it needs and will need from its staff, and from each individual. No doubt changes will be made by both the individual and the organization. There would also be an agreement reached as to how frequently and in what way the review of performance of the goals is to take place. While some will want and need feedback at regular and frequent intervals (say annually), others (a historian, for example) might not need any for several years (although annual progress reports would be filed, but not appraised). When an appraisal does take place, it will be thorough.

We intend to have portfolios used throughout the faculty member's career, not just through the promotion and tenure phases. The fact that individuals are not seriously appraised for the last 25 to 30 years (the modal age for attaining full professor is about 40) of their careers is really an injustice to the faculty member and, most likely, also a loss to the institution. To assume that no new goals will arise once one has acquired the institution's last public recognition until emeritus status, that no assistance is needed to accomplish them, that one can completely take care of oneself, and that to assess a senior professor is an invasion of privacy all fly in the face of reality and the existing evidence. For example, Baldwin and Blackburn (1981) found at least two trying periods for full professors. One is immediately after the final promotion, for it forces a scary look ahead at three decades of the same thing with a security that is formidable. It is too late to try something new but frightfully boring to continue on the same track. The other trying period is the last years before retirement. Faculty in this phase of their careers were often adrift from departmental activities and remote from students, both sources of personal nourishment. In short, senior faculty are in need of a good portfolio process and are entitled to it. In fact, they may need it even more if they are to continue to realize their potential, achieve their aims, and fully contribute to their college or university.

We began this section with a simplifying assumption that faculty goals, not institutional ones, were all that mattered in developing a performance appraisal plan. Now we must consider those cases where faculty goals, while perhaps not incompatible with institutional ones, do not allow the organization to accomplish its functions. Some faculty, for example, may say that they want to spend more time in creative work and less in instruction yet are needed to teach because of student course demand and program needs.

Clearly faculty who were hired to teach will continue to teach to meet the department's obligations. But it may be possible for some to spend less time in instruction if others spend more, or if there are other ways in which the credit hours students demand can be met. If the individual does do more creative work, maybe her or his teaching will improve and an organizational goal will have been achieved. And if that individual does some outstanding work, the status of the unit will be elevated, another institutional goal.

What we are saying is that when portfolios are looked at collectively, some compromises are going to have to be made between the desires of the individual and the needs of the organization. It may turn out that there will be greater differentiation in faculty roles than has been the case. That is, teachers will teach more and researchers will research more. This "two track" system may turn out to be more effective and efficient for many colleges and universities. When staffing imbalances arise and new appointments are sought, those who fill needed roles will know what is expected of them before they sign on.

In general, then, using portfolios as a way of performance assessment is not likely to cause goal achievement problems for the organization. Furthermore, if an organization is having a problem with goal accomplishment, it is more likely to be the result of its having changed (or wishing to change) its mission or image. A regional college wanting to become a national research university is going to ask many of its faculty to become

scholars of high repute. As has been learned, it is not easy to change a nonproducer into a producer. But not having success in changing individual faculty should not be the fault of a performance appraisal plan based on the portfolio process. Just the opposite, in fact, should be true; one would more likely have success in converting individuals through a portfolio process than by a pure dollar reward system for number of pages in print per year, or by one of the other systems we have discussed.

As for merit, individualized portfolio plans would be judged against their individualized goals, not against other's goals. Recognition of accomplishment could take on many forms. Successful growth will have high intrinsic rewards. The institution can recognize the individual with visible honors (titles, publicity releases, a bonus).

The evidence on salary increases is not unequivocal. As stated in Section II, in some instances merit raises reinforced behavior and in others they interfered so that the recommendation was to keep appraisals separate from merit raises. As we indicate in the next section, market conditions need to be considered so it may not be possible for equal merit to receive equal dollar rewards. Faculty may have to increase their trust of others in deciding salaries, since "fairness" systems can rarely be fair.

Merit notwithstanding, we believe the portfolio process will promote individual growth, and the institution will benefit. Both will certainly be in a better condition than they now are when most faculty are competing with one another and many are engaging in activities that are insignificant rather than valuable simply because the system rewards dysfunctional behaviors.

V. The Feasibility of Change

It is one matter to talk about ways in which performance appraisal might take place and to call attention to some factors that colleges and universities need to consider when they undertake performance appraisal of faculty. It is another matter to move from a utopian ideal to a day-to-day reality. The latter requires an assessment of how feasible it is for a college or university to institute an effective new performance appraisal program.

To begin with, administrators need to face the real reasons they are advocating performance appraisal. Today's current interest in performance appraisal is not the first time it has been discussed in higher education. Even though there has not been a documented history of the phenomenon, some truths about its past are worth noting:

1. Calls for performance appraisal of faculty have not received either a constant or a consistent discussion over time.
2. Performance appraisal's sporadic appearance as a "critical issue" seems to be related to *external* pressures for "accountability." When times are tough, "prove your value" say the benefactors in both the private and public sectors. If the pressure becomes intense enough, action is taken; if external pressures are absent or weak, institutions typically do not endorse or practice performance appraisal.
3. The corollary to the second point is that the management's pro-activity in the arena of performance appraisal is minute in contrast to its reactive response. Intuitively or otherwise, in general, administrators have no great faith that performance appraisal will make life better for anyone. The record of failures in institutionalizing performance appraisal is adequate evidence.
4. The evidence also suggests—but does not prove—that institutions that introduce and advocate performance appraisal are either threatened with survival (ergo, get rid of the deadwood faculty) or believe their future depends on a change of mission (e.g., to become research-oriented instead of exclusively teaching- and service-focused, which also means cleaning house and pruning faculty who cannot contribute to the new mission).

Administrators would be wise to make clear the real reasons they are introducing performance appraisal. To tell faculty that performance appraisal is to assist their professional development, when that is not the only aim, is to invite sabotage and to ensure defeat for the plan. Like little kids and dogs who quickly pick up the odor of the phony interloper, faculty quickly detect institutions' ulterior motives. The reasons for implementing performance appraisal have to be justified if the program is to work.

Let us suppose that the above conditions have been met. That is, leadership believes that what we have proposed, or something much like it, a performance appraisal plan incorporating the same research supported principles, will be good for everyone as well as for the institution and wants to have the innovation adopted by the organization. Is it feasible to have such a plan instituted? What needs to be taken into consideration for a sound plan to be accepted and made to work?

To begin with, the portfolio process is fully in accord with college and university norms, values, structures, and practices. In fact, a modified form of the process is used in many institutions for tenure and promotion decisions. On these occasions the person to be evaluated is involved, there is the collecting of evidence and the demonstration of growth, and the judgment by peers is present. Faculty believe the process is right and proper, and that it is fair. So, introducing the innovation will not meet with opposition because it calls for something radically different. That is the good news.

The bad news is that there are many unknowns that can affect the outcome, many of which are outside the institution's control. Instituting a portfolio appraisal process takes time. It is not a quick fix, nor is it without cost. We examine some of the hurdles

we perceive and how adoption might take place in spite of them. We also recognize that conditions will vary from institution to institution.

First, if there is strong conflict between the faculty and the administration, the likelihood of faculty participating in any administration-sponsored plan is essentially zero. Central administration can technically mandate the process, but it will accomplish neither personal nor organizational goals. Faculty will have to see the relative advantage of the process before they commit to it.[7]

Next, for the portfolio appraisal process to be successful, adequate resources must be provided. The programs have real expenses in terms of the number of people involved and the time that is given to the task. Barber's (1983) Education Commission of the States' report on the failure of performance appraisal programs in K-12 districts across the country shows that they were always underfunded when they were instituted. The result was sloppy programs that were easy to attack and have terminated. It should be kept in mind, however, that current practices of complex rating schedules have also been expensive. The portfolio process may not in fact be any more expensive, but we have no way to make realistic estimates of the relative costs of the two methods.

A third ingredient for successfully institutionalizing the innovation will be that faculty are involved in deciding (1) the criteria to be used, (2) the evidence by which performance will be appraised, and (3) the manner in which the evaluation will be conducted. Faculty involvement does not mean that faculty have a right to arbitrarily discard criteria that the administration deems important, but they should have a right to make a case for the appropriateness, and hence the weight, of any criterion advanced. As we saw earlier, different personalities react to assessments of themselves in different ways. How the process is undertaken matters, and that should be taken into consideration by the administration.

Another cost that needs to be considered is for training those who conduct the performance appraisals. Institutions do not ordinarily select department chairs on the basis of the interpersonal skills needed in performance appraisals. In many instances those who the faculty wish as chairs may be untrained or uncomfortable with the appraisal role. The process may have to be led by someone other than the department chair. In all cases, there is a genuine training cost that has to be built into a plan if it is to succeed.

In addition to making clear to the faculty why a performance appraisal program is being introduced, the administrators need to be straightforward about what the anticipated outcomes are. If the plan is to be connected to merit raises, and if faculty have some say in the allocation of their work effort—as well as in the criteria by which it is to be appraised—then they need to be told that equally meritorious work may not always receive the same amount of reward, if that is to be the case. An institution whose first priority is teaching may honor scholarship and give it public recognition but not have it affect salary except in a minor way. A university whose new mission is national recognition may choose to reward publications at a higher rate than good teaching or service. Faculty who allocate their effort toward less-valued activities will receive lower financial reward, regardless of how well their work is accomplished.

In addition, faculty working in institutions that value research highly must recognize that salaries have a market value in part determined by external conditions. Aubrecht (1984, p. 87) nicely draws distinctions between merit, worth, and market value. Whereas physicists and philosophers can have identical salaries in a liberal arts college, those salaries can vary widely in a research university. Hence equally meritorious work may win equal percentage raises but only appreciably fewer dollars for the person with the lower market value.

---

[7] We are following Rogers and Shoemaker's (1971) principles for the successful adoption of an innovation. Seldin (1984, pp.130-133) advances a set of conditions we are advocating here.

Another matter of no small import is the uncertain future which higher education faces. When will enrollments go down, as have been falsely predicted every year now for a decade? When will large numbers of faculty retire and new staff be taken on? Will there be an adequate supply of high quality recruits when they are supposedly needed a decade from now (Bowen & Schuster, 1986)? Many colleges and universities are keeping the number of their full-time, tenured faculty at a minimum to retain flexibility if they have to retrench or change programs.

At the same time, these uncertainties can be used to argue for a portfolio process. The permanent staff is indeed permanent. New long-term hires are still fewer and farther between. Helping existing faculty to continue to grow and develop can be essential for the well-being of the organization. The investment in appraisal with portfolios may be small in relation to the outcomes a college or university needs, as well as what it seeks for individual faculty.

There may be other concerns as well. The Education Commission of the States (Barber, 1983) reports a strong relationship between performance appraisal program failure in K-12 systems with administrative turnover. A new majority on a board or a change in the superintendent can scuttle an operating program simply by devaluing it. The same outcome could occur in a college with a change of the academic dean, if the program was not strongly embedded and operating as an institutional norm with unequivocal endorsement by the president and the board.

If these conditions can be met, the understandings agreed upon, and the resources made available, then a performance appraisal program should work. We suspect the best way to implement it would be to begin with a small unit, a strong department with all of its members performing well, high morale, a healthy relationship with the chair, and clear and strong support from the dean (and all the way to the top). Other units could join in succeeding years. We think simultaneously introducing the innovation throughout an institution, even a small one, will fail.

In conclusion, we believe the portfolio approach to performance appraisal for faculty represents the best mesh of a theoretical systems model with the relevant research findings. While we have advanced an "ideal," we can conceive of modification of that ideal to meet individual or institutional wishes. For example, the portfolio process could also include a uniform rating form, if that seemed absolutely necessary. We recognize that a systematic form keeps the Equal Employment Opportunity Commission at bay, and that may be essential in some institutions, even if the instrument is basically unfair.

Those administrators who want to modify their performance appraisal systems for faculty can draw on some of our specific research findings to do so. For example, a modest financial investment in rater training might well yield positive results not only in the accuracy of the performance appraisal process, but also in faculty support for it.

Finally, to those administrators who are confident that the faculty performance appraisal system they oversee is functioning well for both the institution and individual faculty members, we urge a careful assessment of that conclusion. Here again, a modest investment in some institutional research may be needed to determine the affects that the evaluation is having on the actual improvement of faculty performance and institutional goal accomplishment.

Appendix: Performance Appraisal in Business

While there are many different performance appraisal systems used in business and industry, four general approaches to developing such systems seem to prevail:

1. Comparative procedures
2. Absolute standards
3. Direct indexes
4. Management by objective

Comparative Procedures

In this type of evaluation, an appraisal is made by comparing an employee against others on the criterion of interest. This may be done in four ways:

1. Straight ranking where the evaluator identifies the best performer, second best, and so on to the weakest.
2. Alternative ranking where the evaluator identifies the best and the weakest performers and keeps repeating the process.
3. Paired comparison where one employee at a time is compared with every other.
4. Forced distribution where the evaluator is forced to assign a certain proportion of the unit's employees to each of several performance categories (e.g., above average, average, below average).

Absolute Standards

Systems using absolute standards evaluate individuals against written standards rather than against other employees and measure several factors of global performance rather than a single dimension. Some of the performance appraisal approaches which use absolute standards are:

1. Critical Incidents Technique (CIT). The supervisor observes and notes especially good or bad on-the-job behavior of a subordinate and creates a written record of actual performance to be used as a basis for rating at some time in the future. The general rule is that CIT should be used only for those areas of behavior or performance that cannot be assessed by other more rigorous and systematic methods.

2. Behaviorally Anchored Rating Scales (BARS). These scales are designed to focus on actual behavior and reduce the subjectivity of the rater and diminish interaction problems. BARS attempts a more sophisticated method of developing a rating scale by employing a five-step process: (1) individuals knowledgeable about the job describe specific examples of effective and ineffective job performance; (2) examples are clustered into performance dimensions; (3) a second group of individuals who know the job re-allocate the examples of critical incidents to performance dimensions, with incidents about which there is not substantial agreement as to appropriate assignment dropped (40% to 80% usually considered substantial); (4) second group of participants rates behavior represented by incidents on a seven- to nine-point scale in terms of how effectively the incidents represent performance on that dimension (incidents with widest variability are dropped); (5) a final instrument is developed using a series of scales, one for each performance dimension. The employee's evaluation is determined by summing the scores across all categories. BARS is used when developmental evaluations are the objective.

3. Point-Factor Job Evaluation. This process starts with a detailed job description, usually using four factors and differing their weights—working conditions, problem-solving ability, knowledge required, accountability. Jobs are assessed on

how much of each factor they have, and points are assigned to each factor. A total point score is produced from each evaluation, and the score is translated into a salary level.

4. Job Description Survey (JDS). This instrument identifies and measures the seven job "descriptive constructs": task identity, task significance, skill variety, autonomy, feedback from the work itself, feedback from agents of the work, and opportunity to deal with others.

Direct Indexes

Through the use of direct indexes it is possible to obtain information on performance without engaging in an evaluative process. There are two methods of direct index appraisal.

1. Measures of productivity. These usually quantify some element of output, such as gross sales. In the health sector, measures of occupancy or visits or services produced may be used.
2. Measures of withdrawal. Turnover and absenteeism rates can provide direct information on productivity.

Management by Objective

Management by Objective (MBO) is a goal-setting process that involves managers and their subordinates in jointly establishing goals for work performance and personal development. Thus MBO is more than an evaluation technique: it is a planning, training, and organizational development approach. Managers meet with employees to work out realistic performance objectives; at specified future times the success of the employee in attaining the objectives is evaluated. MBO programs have become increasingly popular in business. The federal government has developed a version of MBO called the Management Excellence Inventory (MEI). Developed by the U.S. Office of Personnel Management (OPM) over a period of four years, MEI was first implemented in mid-1984. It is an instrument designed to assist organizations in identifying managerial job requirements and skills and to help individuals and organizations in assessing management strengths and development needs. The uses of MEI for management development include: clarifying management roles and job requirements; identifying individual strengths and development needs; facilitating individual career planning; making decisions about organizational training and development priorities and resource allocation; planning for management succession; and evaluating training and development programs.

The most frequently voiced criticisms of MBO systems are: (a) an excessive amount of paperwork develops; (b) the process is controlled from the top with little opportunity for real employee participation; and (c) the process turns into a win-lose game between superiors and subordinates.

References

Abramson, P. E., Goldberg, P. A., Greenberg, J. H., & Abramson, U. M. (1978). The talking platypus phenomenon: Competency rating as a function of sex and professional status. *Psychology of Women Quarterly, 2*, 114-124.

Alderfer, C. P. (1969). An empirical test of a new theory of human needs. *Organizational Behavior and Human Performance, 4*, 142-175.

Argyris, C., & Schön, D. A. (1974). *Theory in practice*. San Francisco: Jossey-Bass.

Atkinson, J. W. (1964). *An introduction to motivation*. Princeton, NJ: Van Nostrand.

Aubrecht, J. D. (1984). Better faculty evaluation systems. In P. Seldin (Ed.), *Changing practices in faculty evaluation* (pp. 85-91). San Francisco: Jossey-Bass.

Baldwin, R. G., & Blackburn, R. T. (1981). The academic career as a developmental process: Implications for higher education. *Journal of Higher Education, 52*(6),598-614.

Balzer, W. K. (1986). Biases in the recording of performance related information: The effects of initial impression and centrality of the appraisal task. *Organizational Behavior and Human Decision Processes, 37*(3), 329-347.

Bandura, A. (1977a). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215.

Bandura, A. (1977b). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1980). Gauging the relationship between self-efficacy judgment and action. *Cognitive Therapy and Research, 4*(2), 263-268.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*, 122-147.

Banks, C. G., & Roberson, L. (1985). Performance appraisers as test developers. *Academy of Management Review, 10*(1),128-142.

Barber, L. (1983). *Teacher evaluation and merit pay: A background paper*. Denver: Education Commission of the States. (ERIC Document Reproduction Service No. ED 245 297)

Bare, A. C. (1979). Making growth contracting work. *Journal of the College and University Personnel Association, 30*(4), 15-19.

Bazerman, M. H., Beekun, R. I., & Schoorman, F. D. (1982). Performance evaluation in a dynamic context: A laboratory study of the impact of a prior commitment to the ratee. *Journal of Applied Psychology, 67*(6), 873-876.

Bernardin, H. J., & Beatty, R. W. (1984). *Performance appraisal: Assessing human behavior at work*. Boston: Kent.

Bernardin, H. J., & Pence, E. C. (1980). Effects of rater training: Creating new response sets and decreasing accuracy. *Journal of Applied Psychology, 65*(1), 60-66.

Blackburn, R. T. (1974). The meaning of work in academia. In James I. Doi (Ed.), *Assessing faculty effort* (pp.75-99). San Francisco: Jossey-Bass.

Blackburn, R. T. (1980). Careers for academics and the future production of knowledge. *Annals (American Association for Political and Social Science), 448*, 25-35.

Blackburn, R. T., Boberg, A., O'Connell, C., & Pellino, G. R. (1980). *Project for faculty development program evaluation*. Ann Arbor: University of Michigan, Center for the Study of Higher Education.

Blackburn, R., & Clark, M. J. (1975). An assessment of faculty performance: Some correlates between administrator, colleague, student, and self-ratings. *Sociology of Education, 48*(3), 242-256.

Blackburn, R. T., Horowitz, S. M., Edington, D. W., & Klos, D. M. (1986). University faculty and administrator response to stresses: Correlations with health and job/life satisfactions. *Research in Higher Education, 25*(1), 31-41.

Boal, K. B., & Cummings, L. L. (1981). Cognitive evaluation theory: An experimental test of processes and outcomes. *Organizational Behavior and Human Performance, 28*(3), 289-310.

Borman, W. C. (1979). Individual differences correlates of accuracy in evaluating others' performance effectiveness. *Applied Psychological Measurement, 3*, 103-115.

Borman, W. C. (1983). Implications of personality theory and research for the rating of work performance in organizations. In F. Landy, S. Zedeck, & J. Cleveland (Eds.), *Performance measurement and theory* (pp. 127-165). Hillsdale, NJ: Erlbaum.

Bowen, H. R., & Schuster, J. (1986). *American professors: A national resource imperilled.* New York: Oxford University Press.

Boyer, C. M., & Lewis, D. R. (1985). *And on the seventh day: Faculty consulting and supplemental income.* ASHE-ERIC Higher Education Report No. 3. Washington, DC: Association for the Study of Higher Education.

Burke, R. J., Weitzel, W. J., & Weir, T. (1980). Characteristics of effective interviews of employees' performance review and development: One more time. *Psychological Reports, 47*(3), 683-695.

Calder, B. J., & Staw, B. M. (1975). Self-perception of intrinsic and extrinsic motivation. *Journal of Personality and Social Psychology, 31*(4), 599-605.

Carrell, M. R., & Dittrich, J. E. (1978). Equity theory: The recent literature, methodological considerations, and new directions. *Academy of Management Review, 3*(2), 202-210

Centra, J. A. (1972). *What instructors learn from students: Student and faculty ratings of instruction.* Princeton, NJ: Educational Testing Service.

Centra, J. A. (1975). Colleagues as raters of classroom instruction. *Journal of Higher Education, 46*(1), 327-337.

Cleveland, J. N., & Landy, F. J. (1981). The influence of rater and ratee ages on two performance judgments. *Personnel Psychology, 34*(1), 19-29.

Creswell, J. W. (1985). *Faculty research performance: Lessons from the sciences and the social sciences* (ASHE-ERIC Higher Education Report No. 4). Washington, DC: Association for the Study of Higher Education.

Cummings, L. L. (1983). Performance evaluation systems in the context of individual trust and commitment. In F. Landy, S. Zedeck, & J. Cleveland (Eds.), *Performance measurement and theory* (pp. 89-93). Hillsdale, NJ: Erlbaum.

Deaux, K., & Emswiller, T. (1974). Explanation of successful performance on sex-linked tasks: What is skill for the male, is luck for the female. *Journal of Personality and Social Psychology, 29*(1), 80-85.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology, 18*(1), 105-115.

Deci, E. L. (1972). Intrinsic motivation, extrinsic reinforcement, and inequity. *Journal of Personality and Social Psychology, 22*(1), 113-120.

Deci, E. L. (1975). *Intrinsic motivation.* New York: Plenum.

Deci, E. L. (1976). Notes on the theory and metatheory of intrinsic motivation. *Organizational Behavior and Human Performance, 15*(1), 130-145.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination.* New York: Plenum.

Deets, N. R., & Tyler, D. T. (1986). How Xerox improved its performance appraisals. *Personnel Journal, 65*(4), 50-52.

D'Heilly, P. A. (1975). *Changes in organizational climate, job satisfaction, faculty morale, and attitudes toward educational evaluation related to the introduction of a comprehensive evaluation system.* Unpublished doctoral dissertation, University of Minnesota.

DeNisi, A. S., Randolph, W. A., & Blencoe, A. G. (1983). Potential problems with peer ratings. *Academy of Management Journal, 26*(3), 457-464.

Dipboye, R. L. (1985). Some neglected variables in research on discrimination in appraisals. *Academy of Management Review, 10*(1), 116-127.

Dipboye, R. L., & de Pontbriand, R. (1981). Correlates of employee reactions to performance appraisals and appraisal systems. *Journal of Applied Psychology, 66*(2), 248-251.

Dorfman, P. W., Stephan, W., & Loveland, J. (1986). Performance appraisal behaviors: Supervisor perceptions and subordinate reactions. *Personnel Psychology, 39*(3), 579-597.

Doyle, K., Jr. (1983). *Evaluating teaching.* Lexington, MA: Heath.

Doyle, L. O., Jr., & Webber, P. L. (1978). *Self and student ratings of instruction.* Minneapolis: University of Minnesota, Measurement Services Center.

Ehli, G. J. (1986). Faculty attitudes toward merit pay in South Dakota's public colleges and universities. *Journal of the College and University Personnel Association, 37*(3), 6-12.

Feldman, J. M. (1981). Beyond attribution theory: Cognitive processes in performance appraisal. *Journal of Applied Psychology, 66*(2), 127-148.

Feldman, J. M. (1983). Practice, practicality, and prospects of training for performance appraisal. In C. Banks & K. Murphy (Chairs), *Performance appraisal research applied to real organizations: Narrowing the gap.* Symposium conducted at the 91st meeting of the American Psychological Association, Anaheim, CA.

Finkelstein, M. J. (1984). *The American academic profession: A synthesis of social scientific inquiry since World War II.* Columbus: Ohio State University Press.

Finn, R. H., & Fontaine, P. A. (1984). Performance appraisal: Some dynamics and dilemmas. *Public Personnel Management, 13*(3), 335-343.

French, J. R. P., Caplan, R. D., & Harrison, R. V. (1982). *The mechanisms of job stress and strain.* Harrison, NY: Wiley.

Fulk, J., Brief, A. P., & Barr, S. H. (1985). Trust-in-supervisor and perceived fairness and accuracy of performance evaluations. *Journal of Business Research, 13*(4), 301-313.

Gaff, J. (no date). *PIRIT project.* Mimeographed.

Goodman, P. S. (1977). Social comparison process in organizations. In B. M. Staw & G. R. Salancik (Eds.), *New directions in organizational behavior* (Vol. 1, pp. 97-132). Chicago, IL: St. Clair Press.

Hackman, J. R., & Oldham, G. R. (1976). Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance, 16*(2), 250-279.

Hamilton, D. L., Katz, L. B., & Leirer, V. O. (1980). Organizational processes in impression formation. In R. Hastie, T. Ostrom, R. Ebbesen, R. Wyer, D. Hamilton, & D. Carlston, (Eds.), *Person memory: The cognitive basis of social perception* (pp. 120-150). Hillsdale, NJ: Erlbaum.

Hamner, W. C., Kim, J. S., Baird, L., & Bigoness, W. J. (1974). Race and sex as determinants of ratings by potential employers in a simulated work sampling task. *Journal of Applied Psychology, 59*(6), 705-711.

Harlan, A., Kerr, J., & Kerr, S. (1977). Preference for motivator and hygiene factors in a hypothetical interview situation: Further findings and some implications for the employment interview. *Personnel Psychology, 30*(4), 555-566.

Heilman, M. E., & Stopeck, M. H. (1985). Being attractive, advantage or disadvantage? Performance-based evaluations and recommended personnel actions as a function of appearance, sex, and job type. *Organizational Behavior and Human Decision Processes, 35*(2), 205-215.

Hind, R. R., Dornbusch, S. M., & Scott, W. R. (1974, Winter). A theory of evaluation applied to a university faculty. *Sociology of Education, 47*, 114-128.

Hobson, C. J., Mendel, R. M., & Gibson, F. W. (1981). Clarifying performance appraisal criteria. *Organizational Behavior and Human Performance, 28*(2), 164-188.

Hogan, E. A. (1987). Effects of prior expectations on performance ratings: A longitudinal study. *Academy of Management Journal, 30*(2), 354-368.

Holley, W., Jr., Halpin, G., & Johnson, G. (1982). Assessment of faculty performance appraisal. *Journal of the College and University Personnel Association, 33*(1), 14-18.

Holzbach, R. (1978). Rater bias in performance ratings: Superior,self and peer ratings. *Journal of Applied Psychology, 63*(5), 579-588.

Ilgen, D. R., & Favero, J. L. (1985). Limits in generalization from psychological research to performance appraisal processes. *Academy of Management Review, 10*(2), 311-321.

Ilgen, D. R., & Feldman, J. M. (1983). Performance appraisal: A process focus. In L. L. Cummings & B. M. Staw (Eds.), *Research in organizational behavior* (Vol. 5, pp. 141-197). Greenwich, CT: JAI Press.

Ilgen, D. R., Peterson, R. B., Martin, B. A., & Boeschen, D. A. (1981). Supervisor and subordinate reactions to performance appraisal sessions. *Organizational Behavior and Human Performance, 28*(3), 311-330.

Ivancevich, J. M., & Matteson, M. T. (1983). Stress and performance. In R. M. Steers & L. W. Porter (Eds.), *Motivation and work behavior* (3rd ed., pp. 375-385). New York: McGraw-Hill.

Jacobs, R., Kafry, D., & Zedeck, S. (1980). Expectations of behaviorally anchored rating scales. *Personnel Psychology, 33*(3), 595-640.

Kasten, K. L. (1984, July/August). Tenure and merit pay as rewards for research, teaching, and service at a research university. *Journal of Higher Education, 55*(4), 500-514.

Korman, A. K. (1971). Expectancies as determinants of performance. *Journal of Applied Psychology, 55*(3), 218-222.

Ladd, E. C., Jr. (1979). The work experience of American college professors: Some data and an argument. In *Current Issues in Higher Education* (pp. 3-12). Washington, DC: American Association for Higher Education.

Landy, F. J., Barnes, J. L., & Murphy, K. R. (1978). Correlates of perceived fairness and accuracy of performance evaluation. *Journal of Applied Psychology, 63*(6), 751-754.

Landy, F. J., Barnes-Farrell, J., & Cleveland, J. N. (1980). Perceived fairness and accuracy of performance evaluation: A follow-up. *Journal of Applied Psychology, 65*(3), 355-356.

Landy, F. J., & Becker, W. S. (1987). Motivation theory reconsidered. In L. L. Cummings & B. M. Staw (Eds.), *Research in organizational behavior* (Vol. 9, pp. 1-40). Greenwich, CT: JAI Press.

Landy, F., & Farr, J. (1980). Performance ratings. *Psychological Bulletin, 8*(1), 72-107.

Landy, F., Zedeck, S., & Cleveland, J. (1983). *Performance measurement and theory*. Hillsdale, NJ: Erlbaum.

Lawrence, J. H. (1988). Faculty motivation and the teaching task: Some old and new conceptualizations. In J. S. Stark & L. A. Mets (Eds.), *Studying learning and the academic environment*, New Directions for Institutional Research no. 57. San Francisco: Jossey-Bass.

Locke, E. A. (1968). Toward a theory of task motivation and incentives. *Organizational Behavior and Human Performance, 3*(2), 157-189.

Locke, E. A. (1976). The nature and causes of job satisfaction. In M. D. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 1297-1349). Chicago: Rand McNally.

Locke, E. A. (1978). The ubiquity of the techniques of goal setting in theories and approaches to employee motivation. *Academy of Management Review, 3*(4), 594-601.

Love, K. G. (1981). Comparison of peer assessment methods: Reliability, validity, friendship bias, and user reaction. *Journal of Applied Psychology, 66*(4), 451-457.

Maehr, M. L., & Braskamp, L. A. (1986). *The motivation factor: A theory of personal investment*. Lexington, MA: Heath.

Marsh, H. W., Overall, J. U., & Kessler, S. P. (1979). Validity of student evaluations of instructional effectiveness: A comparison of faculty self evaluations and evaluations by their students. *Journal of Educational Psychology, 71*, 50.

Maruyama, G., & Miller, N. (1980). Physical attractiveness, race, and essay evaluation. *Personality and Social Psychology Bulletin, 6*, 384-390.

Maslow, A. H. (1954). *Motivation and personality*. New York: Harper.

McClelland, D. C. (1961). *The achieving society*. Princeton, NJ: Van Nostrand.

McClelland, D., Atkinson, J. W., Clark, J. W., & Lowell, E. L. (1953). *The achievement motive*. New York: Appleton-Century-Crofts.

McGregor, D. (1957). An uneasy look at performance appraisal. *Harvard Business Review, 35*(3), 89-94.

McGregor, D. (1967). *The professional manager*. New York: McGraw Hill.

McKeachie, W. J., Pintrich, P. R., Lin, Y., & Smith, D. A. F. (1986). *Teaching and learning in the college classroom: A review of the research literature* (Tech. Rep. No. 86-B-001). Ann Arbor, MI: University of Michigan, National Center for Research to Improve Postsecondary Teaching and Learning.

Meyer, H. H. (1975). The pay for performance dilemma. *Organizational Dynamics, 3*, 39-50.

Meyer, H. H. (1980). Self-appraisal of job performance. *Personnel Psychology, 33*(2), 291-293.

Miller, J. R., & Young, J. I. (1979). Merit pay: An unexamined concept in higher education. *Journal of the College and University Personnel Association, 30*(4), 7-14.

Miller, R. I. (1974). *Evaluating faculty performance*. San Francisco: Jossey-Bass.

Miller, R. I. (1987). *Evaluating faculty for promotion and tenure*. San Francisco: Jossey-Bass.

Mitchell, T. R. (1982). Motivation: New directions for theory, research, and practice. *Academy of Management Review, 7*(1), 80-88.

Morris, J., & Steers, R. M. (1980). Structural influences on organizational commitment. *Journal of Vocational Behavior, 17*(1), 50-57.

Mortimer, R. P., Bagshaw, M., & Masland, A. T. (1985). *Flexibility in academic staffing: Effective policies and practices* (ASHE-ERIC Higher Education Report No. 1). Washington, DC: Association for the Study of Higher Education.

Mount, M. K. (1983). Comparisons of managerial and employee satisfaction with a performance appraisal system. *Personnel Psychology, 36*(1), 99-110.

Mount, M. K. (1984). Supervisor, self, and subordinate ratings of performance and satisfaction with supervisor. *Journal of Management, 10*(3), 305-320.

Mowday, R. T., Porter, L., W., & Steers, R. M. (1982). *Employee-organizational linkages.* New York: Academic Press.

Murray, H. A. (1938). *Explorations in personality.* New York: Oxford University Press.

Nadler, D., & Lawler, E. E., III. (1977). Motivation: A diagnostic approach. In J. R. Hackman, E. E. Lawler III, & L. Porter (Eds.), *Perspectives on behavior in organizations* (pp. 26-38). New York: McGraw Hill.

Nisbett, R. D., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology, 35*(4), 250-256.

Ormrod, J. E. (1986). Predictors of faculty dissatisfaction with an annual performance evaluation. *Journal of the College and University Personnel Association, 37*(3), 13-18.

Ory, J. C., & Braskamp, L. A. (1981). Faculty perceptions of the quality and usefulness of three types of evaluative information. *Research in Higher Education, 15*(3), 271-282

Patz, A. (1975). Performance appraisal: Useful but still resisted. *Harvard Business Review, 53*(3), 74-80.

Pearce, J. L., & Porter, L. W. (1986). Employee responses to formal performance appraisal feedback. *Journal of Applied Psychology, 71*(2), 211-218.

Pearce, J. L., Stevenson, W., & Perry, J. L. (1985). Managerial compensation based on organizational performance: A time series analysis of the effects of merit pay. *Academy of Management Journal, 28*(2), 261-278.

Perry, J. L., & Porter, L. W. (1982). Factors affecting the context for motivation in public organizations. *Academy of Management Review, 7*(1), 89-98.

Pinder, C. C. (1976). Addivity versus nonaddivity of intrinsic and extrinsic incentives: Implications for work motivation, performance, and attitudes. *Journal of Applied Psychology, 61*(6), 693-700.

Poole, L. H., & Dellow, D. A. (1983). Evaluation of full-time faculty. In A. Smith (Ed.), *Evaluating faculty and staff.* New Directions for Community Colleges, no. 41. San Francisco: Jossey-Bass.

Porter, L. W., & Lawler, E. E., III. (1968). *Managerial attitudes and performance.* Homewood, IL: Irwin.

Porter, L. W., & Miles, R. P. (1974). Motivation and management. In J. W. McGuire (Ed.), *Contemporary management: Issues and viewpoints.* Englewood Cliffs, NJ: Prentice-Hall.

Prince, J. B., & Lawler, E. E., III. (1986). Does salary discussion hurt the developmental performance appraisal? *Organizational Behavior and Human Decision Processes, 37*(3), 357-375.

Pulakos, E. (1984). *The interactive effects of format and training on performance appraisal accuracy.* Unpublished doctoral dissertation, Michigan State University, East Lansing.

Raelin, J. A. (1985). The basis of the professional's resistance to managerial control. *Human Resource Management, 24*(2), 147-175.

Raynor, J. E., & Entin, E. E. (1982). *Motivation, career striving and aging.* New York: Hemisphere Publishing Corporation.

Renz, F. J. (1984). *Study examining the issues of faculty evaluation.* Austin: University of Texas. (ERIC Document Reproduction Service No. ED 243 559)

Rogers, E. M., & Shoemaker, F. F. (1971). *Communication of innovations.* New York: Free Press.

Rose, G. L. (1978). Sex effects of effort attributions in managerial performance evaluation. *Organizational Behavior and Human Performance, 21*(3), 367-378.

Rosen, B., Jerdee, T. H., & Lunn, R. O. (1981). Effects of performance appraisal format, age, and performance level on retirement decisions. *Journal of Applied Psychology, 66*(4), 515-519.

Rossi, A. S. (1980). Life-span theories and women's lives. *Signs, 6*(1), 4-32.

Schmitt, N., & Lappin, M. (1980). Race and sex as determinants of the mean and variance of performance ratings. *Journal of Applied Psychology, 65*(4), 428-435.

Schneider, D. J. (1973). Implicit personality theory: A review. *Psychological Bulletin, 79*(2), 294-309.

Schwab, D. P., & Heneman, H. H. (1978). *Age stereotyping in performance appraisal.* Paper presented at the National Academy of Management, San Francisco.

Seldin, P. (1984). *Changing practices in faculty evaluation.* San Francisco: Jossey-Bass.

Shapiro, G., & Dessler, G. (1985). Are self-appraisals more realistic among professionals or non-professionals in health care? *Public Personnel Management, 14*(3), 285-291.

Shields, W. S. (1984). *A systems approach to performance appraisal system development in a large organization.* Unpublished doctoral dissertation, York University, Downsview, Canada.

Shrauger, J. S., & Schoeneman, T. S. (1979). Symbolic interactionist view of self concept: Through the looking glass darkly. *Psychological Bulletin, 86*(3), 549-573.

Skinner, B. F. (1971). *Beyond freedom and dignity.* New York: Knopf.

Staw, B. M. (1976). Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance, 16*(1), 27-44.

Staw, B. M. (1981). The escalation of commitment to a course of action. *Academy of Management Review, 6*(4), 577-587.

Staw, B. M. (1984). Motivation research versus the art of faculty management. In J. L. Bess (Ed.), *College and university organization: Insights from the behavioral sciences* (pp. 63-83). New York: New York University Press.

Steel, B. S. (1985). Participative performance appraisal in Washington: An assessment of post-implementation receptivity. *Public Personnel Management, 14*(2), 153-171.

Steers, R. M. (1977). Antecedents and outcomes of organizational commitment. *Administrative Science Quarterly, 22*(1), 46-56.

Steers, R. M. (1981). *Introduction to organization behavior.* Glenview, IL: Scott, Foresman.

Steers, R. M. (1983). Murray's manifest needs theory. In R. M. Steers & L. W. Porter (Eds.), *Motivation and work behavior* (pp. 42-49). New York: McGraw-Hill.

Steers, R. M., & Lee, T. W. (1983). Facilitating effective performance appraisals: The role of employee commitment and organizational climate. In F. Landy, S. Zedeck, & J. Cleveland (Eds.), *Performance measurement and theory* (pp. 75-88). Hillsdale, NJ: Erlbaum.

Steers, R. M., & Porter, L. W. (1974). The role of task-goal attributes in employee performance. *Psychological Bulletin, 81*(3), 434-452.

Steers, R. M., & Porter, L. W. (Eds.). (1983). *Motivation and work behavior.* New York: McGraw Hill.

Stone, E. F., & Stone, D. L. (1984). The effects of multiple sources of performance feedback and feedback favorability in self perceived task competence and perceived feedback accuracy. *Journal of Management, 10*(3), 371-378.

Taylor, M. S., Fisher, C. D., & Ilgen, D. R. (1984). Individuals' reactions to performance feedback in organizations: A control theory perspective. In R. M. Rowland & G. R. Ferris (Eds.), *Research in personnel and human services management* (pp.81-124). Greenwich, CT: JAI Press.

Taylor, S. E., & Fiske, S. T. (1978). Salience, attention, and attributions: Top of the head phenomena. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. II, pp. 250-288). New York: Academic Press.

Terpstra, D. E., Olson, P. D., & Lockeman, B. (1982). The effects of MBO on levels of performance and satisfaction among university faculty. *Group & Organization Studies, 7*(3), 353-366.

Thornton, G. (1980). Psychometric properties of self appraisals of job performance. *Personnel Psychology, 33*(2), 263-271.

Tucker, A. (1984). *Chairing the academic department.* New York: Macmillan.

Vroom, V. (1964). *Work and motivation.* New York: Wiley.

Wahba, M. A., & Bridwell, L. G. (1973). Maslow reconsidered: A review of research on the need hierarchy theory. In R. M. Steers & L. W. Porter (Eds.), *Motivation and work behavior* (pp. 34-41). New York: McGraw Hill.

Watts, L. (1985). *Assessment of an organizational control system.* Tempe, AZ. (ERIC Document Reproduction Service No. ED 258 510)

Weiner, B. (1972). *Theories of motivation: from mechanism to cognition.* Chicago: Rand McNally.

Wexley, K. N., & Pulakos, E. D. (1982). Sex effects on performance ratings in manager-subordinate dyads: A field study.*Journal of Applied Psychology, 67*(4), 433-430.

Wexley, K. N., Singh, J. P., & Yukl, G. A. (1973). Subordinate personality as a moderator of the effects of participation in three types of appraisal interviews. *Journal of Applied Psychology, 58*(1), 54-59.

Williams, K. J., De Nisi, A., Blencoe, A., & Cafferty, T. (1985). The role of appraisal purpose: Effects of purpose on information acquisition and utilization. *Organizational Behavior and Human Decision Processes, 35*(3), 314-339.

Zammuto, R., London, M., & Rowland, K. (1982). Organization and rater differences in performance appraisals. *Personnel Psychology, 35*(2), 643-658.

Zedeck, S., & Cascio, W. (1982). Performance appraisal decisions as a function of rater training and purpose of the appraisal. *Journal of Applied Psychology, 67*(6), 752-758.

Zey-Farrell, M., & Ervin, D. (1985). Achieving congruent actions and intentions: An empirical assessment of faculty work in a regional public university. *Research in Higher Education, 22*(4), 347-369.

Bibliography

Atkinson, J. W. (1983). *Personality, motivation, and action.* New York: Praeger.

Bailyn, L. (1985). Autonomy in the industrial R & D lab. *Human Resource Management, 24*(2), 129-146.

Balcazar, F. E., Hopkins, B. L., & Suarez, Y. (1985-86). A critical objective review of performance feedback. *Journal of Organizational Behavior Management, 7*(3-4), 65-89.

Berkowitz, E. N. (1980). Role theory, attitudinal constructs, and actual performance: A measurement issue. *Journal of Applied Psychology, 65*(2), 240-245.

Bland, W. D. (1982-83). Role orientations and evaluation of professional performance. *Community College Review, 10*(3), 27-32.

Brown, W. S. (1982). Merit pay in the university environment. *Journal of the College and University Personnel Association, 33*(1), 27-31.

Brown, W. S. (1983). Pay for performance: The merit pay concept in an academic environment. *Journal of the College and University Personnel Association, 34*(3), 23-26.

Brown, W. S. (1984). Performance review instruments and merit pay programs in an academic environment. *Journal of the College and University Personnel Association, 35*(1), 7-14.

Burke, R. J., Deszca, G., & Weitzel, W. (1982). Subordinate expectations and satisfaction with the performance appraisal interview. *Journal of Psychology, III*(1), 41-49.

Carrell, M. R. (1978). A longitudinal field assessment of employee perceptions of equitable treatment. *Organizational Behavior and Human Performance, 21*(1), 108-118.

Carroll, J. G., & Goldberg, S. R. (1975). *Igniting the vital spark: A process of peer review and consultation in postsecondary education.* Chicago: American Educational Research Association. (ERIC Document Reproduction Service No. ED 264 766)

Carver, F. D., & Sergiovanni, T. J. (1969). *Organizations and human behavior: Focus on schools.* New York: McGraw-Hill.

Cascio, W. J. & Ramos, R. A. (1986). Development and application of a new method for assessing job performance in behavioral/economic terms. *Journal of Applied Psychology, 71*(1), 20-28.

Catania, J. C. (1977). *Faculty evaluation within collective bargaining constraints.* Paper presented at the annual meeting of the Association of Community and Junior Colleges, Denver. (ERIC Document Reproduction Service No. ED 142 251)

Cohen, P. A. (1981). Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies. *Review of Educational Research, 51*(3), 281-309.

DeCotiis, T. A., & Dyer, L. (1978). Implication of the age discrimination amendments for performance appraisal. *Journal of the College and University Personnel Association, 29*(4), 21-24.

DeNisi, A. S., Cafferty, T. P., & Meglino, B. M. (1984). A cognitive view of the performance appraisal process: A model and research proposition. *Organizational Behavior and Human Performance, 33*(3), 360-396.

Duncan, P. K., & Bruwelheide, L. R. (1985-86). Feedback use and possible behavioral functions. *Journal of Organizational Behavior Management, 7*(3/4), 91-114.

Fisher, C. D. (1978). The effects of personal control, competence, and extrinsic reward systems on intrinsic motivation. *Organizational Behavior and Human Performance, 21*(3), 273-288.

Frankle, A. H. (1978). Sequential response shift rate: A correlate of human adaptivity measurable with existing personality inventories. *Journal of Psychology, 98*(1), 129-143.

Gaff, J. G., & Wilson, R. C. (1975). Faculty impact on students. In R. C. Wilson, J. Gaff, E. Dienst, L. Wood, & J. Bavry (Eds.), *College professors and their impact on students*. New York: Wiley.

Gioia, D. A., & Sims, H. P., Jr. (1986). Cognition-behavior connections: Attribution and verbal behavior in leader-subordinate interactions. *Organizational Behavior and Human Decision Processes, 37*(2), 197-229.

Greller, M. M. (1980). Evaluation of feedback sources as a function of role and organizational level. *Journal of Applied Psychology, 65*(1), 24-27.

Gunn, B. (1982). Evaluating faculty performance: A holistic approach. *Journal of the College and University Personnel Association, 33*(4), 23-30.

Hanser, L. M., & Muchinsky, P. M. (1978). Work as an information environment. *Organizational Behavior and Human Performance, 21*(1), 47-60.

Hellriegel, D., Slocum, J. W., & Woodman, R. W. (1983). *Organizational Behavior* (3rd ed). New York: West.

Herzberg, F., Mausner, B., & Snyderman, B. (1959). *The motivation to work*. New York: Wiley.

Hom, P. W., De Nisi, A., Kinicki, A., & Bannister, B. (1982). Effectiveness of performance feedback from behaviorally anchored rating scales. *Journal of Applied Psychology, 67*(5), 568-576.

Huber, V. L. (1985-86). The interplay of goals and promises of pay-for-performance on individuals and group performance: An operant interpretation. *Journal of Organizational Behavior Management, 7*(3/4), 45-63.

Hunter, J. E., & Schmidt, F. L. (1983). Quantifying the effects of psychological interventions on employee job performance and work-force productivity. *American Psychologist, 38*(4), 473-478.

Ivancevich, J. M. (1980). A longitudinal study of behavioral expectation scales: Attitudes and performance. *Journal of Applied Psychology, 65*(2), 139-146.

Kim, J. S. (1984). Effect of behavior plus outcome goal setting and feedback on employee satisfaction and performance. *Academy of Management Journal, 27*(1), 139-149.

Ladd, E. C., & Lipset, S. M. (1976). *Survey of the social, political and educational perspectives of American college and university faculty* (Vols. 1-2). Storrs, CT: University of Connecticut.

Larson, J. R., Jr. (1984). The performance feedback process: A preliminary model. *Organizational Behavior and Human Performance, 33*(1), 42-76.

Larson, J. R., Jr. (1986). Supervisors' performance feedback to subordinates: The impact of subordinate performance valence and outcome dependence. *Organizational Behavior and Human Decision Processes, 37*(3), 391-408.

Licata, C. M. (1986). *Post-tenure faculty evaluations: Threat or opportunity*. ASHE-ERIC Higher Education Report No. 1. Washington, DC: Association for the Study of Higher Education.

Liden, R. C., & Mitchell, T. R. (1985). Reactions to feedback: The role of attributions. *Academy of Management Journal, 28*(2), 291-308.

Maehr, M. L. (1974). *Sociocultural origins of achievement*. Monterey, CA: Brooks/Cole.

McClelland, D. C. (1985). *Human motivation*. Glenview, IL: Scott, Foresman.

McIntosh, T. H., & Van Koevering, T. E. (1986). Six-year case study of faculty peer reviews, merit ratings, and pay awards in a multi-disciplinary department. *Journal of the College and University Personnel Association, 37*(1), 5-14.

Mitchell, T., Green, S., & Wood, R. (1981). An attribution model of leadership and the poor performing subordinate. In L. Cummings & B. Staw (Eds.), *Research in organizational behavior* (Vol. 3, pp. 197-234). Greenwich, CT: JAI Press.

Newton, R. R. (1982). Performance evaluation in education. *Journal of the College and University Personnel Association, 33*(2), 39-43.

Phillips, J. S., & Lord, R. G. (1980). Determinants of intrinsic motivation: Locus of control and competence information as components of Deci's cognitive evaluation theory. *Journal of Applied Psychology, 65*(2), 211-218.

Reisman, B. (1986). Performance evaluation for tenured faculty: Issues and research. *Liberal Education, 72*(1), 73-87.

Ronen, S. (1978). Personal values: A basis for work motivational set and work attitude. *Organizational Behavior and Human Performance, 21*(1), 80-107.

Root, L. R. (1987). Faculty evaluation: Reliability of peer assessments of research, teaching, and service. *Research in Higher Education, 26*(1), 71-84.

Silverman, S. B., & Wexley, K. N. (1984). Reaction of employees to performance appraisal interviews as a function of their participation in rating scale development. *Personnel Psychology, 37*(4), 703-710.

Simon, K. M. (1979). Self-evaluative reactions: The role of personal valuation of the activity. *Cognitive Therapy and Research, 3*(1), 111-116.

Smith, A. (Ed.). (1983). *Evaluating faculty and staff.* New Directions for Community Colleges, no. 41. San Francisco: Jossey-Bass.

Spencer, D. G., & Steers, R. M. (1981). Performance as a moderator of the job satisfaction-turnover relationship. *Journal of Applied Psychology, 66*(4), 511-514.

Stein, J. M. (1986). Public employee productivity: Can outcomes be validly measured at the jurisdictional level? *Public Personnel Management, 15*(2), 111-117.

Stone, D. L., Gueutal, H. G., & McIntosh, B. (1984). The effects of feedback sequence and expertise of the rater on perceived feedback accuracy. *Personnel Psychology, 37*(3), 487-506.

Stone, T. H. (1973). An examination of six prevalent assumptions concerning performance appraisal. *Public Personnel Management, 2*(6), 408-414.

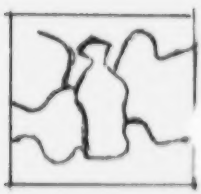
Von Glinow, M. A. (1985). Reward strategies for attracting, evaluating, and retaining professionals. *Human Resource Management, 24*(2), 191-206.

Weiner, B. (1980). *Human motivation.* New York: Holt, Rinehart & Winston.

Wexley, K. N., & Pulakos, E. D. (1983). The effects of perceptual congruence and sex on subordinates' performance appraisals of their managers. *Academy of Management Journal, 26*(4), 666-676.

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review, 66*(3), 297-333.

Young, R. J., & Gwalamubisi, Y. (1986). Perceptions about current and ideal methods and purposes of faculty evaluation. *Community College Review, 13*(4), 27-33.



NATIONAL CENTER FOR
RESEARCH TO IMPROVE
POSTSECONDARY
TEACHING and LEARNING